

TECHNISCHE UNIVERSITÄT
CHEMNITZ

FAKULTÄT FÜR WIRTSCHAFTSWISSENSCHAFTEN

Ergebnisheterogenität in der Strategischen Managementfor- schung – Analyse und Lösungsmöglichkeiten

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. pol.)

an der Wirtschaftswissenschaftlichen Fakultät
der Technischen Universität Chemnitz

Vorgelegt von:

Dipl.-Kfm. Norbert Steigenberger

Beckerstraße 26 c

09120 Chemnitz

Gutachter:

Prof. Dr. Peter Pawlowsky, Technische Universität Chemnitz

Prof. Dr. Dieter Wagner, Universität Potsdam

Chemnitz, 14. September 2010

Überblick

1. Einleitung	1
1.1 Problemstellung	1
1.2 Forschungsfragen, zentrales Erkenntnisinteresse, Aufbau und Einordnung	5
2. Grundlagen.....	8
2.1 Die Rolle von Theorie, Empirie und Hypothese in der Wissensgenerierung	8
2.2 Der Leistungsbegriff	19
2.3 Strategische Managementforschung	25
2.4 Zwischenfazit.....	30
3. Theoriegeleitete Analyse.....	33
3.1 Überblick	33
3.2 Das Problem der Populationsselektion	34
3.3 Das Appropriationsproblem	44
3.4 Das Time-Lag Problem.....	70
4 Qualitative Literaturreview	77
4.1 Forschungsansatz	77
4.2 Ergebnisse	89
4.3 Diskussion.....	96
5. Metaanalyse.....	99
5.1 Forschungsansatz	99
5.2 Hypothesenentwicklung.....	103
5.3 Daten und Operationalisierung.....	106
5.4 Analyse	113
5.5 Diskussion der Ergebnisse	141
5.6 Abgrenzung und Limitationen.....	153
6. Zusammenführung und Diskussion	157
6.1 Zusammenfassung der Kernergebnisse	157
6.2 Diskussion und Ausblick	165
Anhänge.....	174

Inhaltsverzeichnis

1. Einleitung	1
1.1 Problemstellung	1
1.2 Forschungsfragen, zentrales Erkenntnisinteresse, Aufbau und Einordnung	5
2. Grundlagen.....	8
2.1 Die Rolle von Theorie, Empirie und Hypothese in der Wissensgenerierung	8
2.1.1 Induktion, Deduktion und das Problem der Verallgemeinerbarkeit	8
2.1.2 Wissensfortschritt durch Falsifikation	12
2.1.3 Hypothesen und Hypothesentests	15
2.1.4 Operationalisierung und Indikatoren	17
2.2 Der Leistungsbegriff	19
2.2.1 Leistung – Das Definitionsproblem	19
2.2.2 Business Performance Measurement – Abgrenzung.....	23
2.3 Strategische Managementforschung	25
2.4 Zwischenfazit.....	30
3. Theoriegeleitete Analyse.....	33
3.1 Überblick	33
3.2 Das Problem der Populationsselektion	34
3.2.1 Populationsselektion – Die Problemlage	34
3.2.2 Populationsselektion – wissenschaftstheoretisches Problem	37
3.2.3 Populationsselektion – Literaturüberblick.....	38
3.2.4 Diskussion	43
3.3 Das Appropriationsproblem	44
3.3.1 Organisationstheoretische Grundlagen	45
3.3.2 Wertschöpfung und ihre Verteilung auf Stakeholder-Gruppen	51
3.3.3 Auswirkungen auf die Leistungsoperationalisierung: Verhandlungsmacht als Moderatorvariable.....	59
3.3.4 Die Appropriationstheorie und die unterschiedlichen Indikatoren.....	60
3.3.5 Das Konzept der Kernperformance und ihre mathematische Modellierung.....	65
3.4 Das Time-Lag Problem.....	70
3.4.1 Die Problemlage, Hypothesenentwicklung.....	70
3.4.2 Lösungsansätze	73

4 Qualitative Literaturreview	77
4.1 Forschungsansatz	77
4.1.1 Fragestellung	77
4.1.2 Qualitative Literaturreview – Methodik	78
4.1.3 Datenquellen und Operationalisierung	79
4.1.4 Limitationen.....	88
4.2 Ergebnisse	89
4.2.1 Leistungsmessung in der Strategischen Managementforschung (Forschungsfrage F1).....	89
4.2.2 Findet sich ein dominanter Leistungsindikator (Forschungsfrage F2)?	91
4.2.3 Der Umgang mit dem Problem der Verallgemeinerbarkeit (Forschungsfrage 3)	92
4.2.4 Der Umgang mit dem Time-Lag Problem	94
4.3 Diskussion.....	96
5. Metaanalyse.....	99
5.1 Forschungsansatz	99
5.1.1 Fragestellung	99
5.1.2 Metaanalyse – Methodik.....	99
5.2 Hypothesenentwicklung.....	103
5.3 Daten und Operationalisierung.....	106
5.3.1 Datensatz	106
5.3.2 Operationalisierung	109
5.3.3 Synopse.....	111
5.4 Analyse	113
5.4.1 Statistische Korrekturfaktoren und Vorüberlegungen	113
5.4.2 Explorative Analyse.....	117
5.4.3 Hypothesentests.....	118
5.4.4 Weiterführende Untersuchungen	134
5.5 Diskussion der Ergebnisse	141
5.5.1 Populationsselektion	142
5.5.2 Appropriation	146
5.5.3 Time-Lag	152
5.6 Abgrenzung und Limitationen.....	153
5.6.1 Limitationen.....	153
5.6.2 Einordnung in den wissenschaftlichen Diskurs.....	155

6. Zusammenführung und Diskussion	157
6.1 Zusammenfassung der Kernergebnisse	157
6.1.1 Erkenntnisinteresse, Forschungsbedarf und Gang der Arbeit.....	157
6.1.2 Epistemologische Betrachtungen, der Leistungsbegriff und das Fachgebiet der Strategischen Managementforschung.....	157
6.1.3 Konzeptionelle Ansätze zur Erklärung der Ergebnisheterogenität in der Strategischen Managementforschung	159
6.1.4 Empirische Analyse	162
6.2 Diskussion und Ausblick	165
6.2.1 Diskussion und Implikationen für die wissenschaftliche Praxis.....	165
6.2.2 Grenzen dieser Untersuchung und weiterer Forschungsbedarf.....	172
Anhänge.....	174
A.1 Verzeichnisse	174
Abbildungsverzeichnis	174
Tabellenverzeichnis	174
Abkürzungsverzeichnis	176
Literaturverzeichnisse.....	177
A.2 Tabellen.....	214
A.3 Inhaltlicher Anhang.....	217
A.3.1 Regionaler Fokus der Strategischen Managementforschung.....	217
A.3.2 Weitere Methodenprobleme in der Strategischen Managementforschung.....	219

1. Einleitung

1.1 Problemstellung

Eine Grunderkenntnis der strategischen Managementforschung besagt, dass sich die Leistung von Unternehmen nicht nur kurzfristig unterscheidet, sondern dass es Unternehmen gibt, die in der Lage sind, dauerhaft überdurchschnittliche Ergebnisse zu erzielen und damit Überprofite zu erwirtschaften (Buo & Satorra, 2007). Eine Beobachtung, die mit dem ökonomischen (neoklassischen) Standardmodell nicht erklärbar ist und die von Rumelt und Kollegen treffend als *problem of persistent profits* bezeichnet wurde (Rumelt, Schendel & Teece, 1991). Dem Standardmodell folgend würden Überprofite eines Unternehmens zu Imitation durch andere Marktteilnehmer und Überprofite in einem Markt zu Marktzutritten durch weitere Unternehmen führen, der Konkurrenzdruck, bzw. die Erhöhung des Angebots, würde zu Preis Anpassungen nach unten führen bzw. die Imitation die Kostenvorteile des Überprofite generierenden Unternehmens aufzehren (Schöler, 2004). Auch wenn das Ausmaß der Performance-Unterschiede und ihre Dauerhaftigkeit nicht unumstritten sind (vergleiche Wiggins & Ruefli, 2002; Powell & Arregle, 2007; Buo & Satorra, 2007), stellt ihre Existenz dennoch die Grundfrage der Strategischen Managementforschung dar und schafft Erklärungsbedarf: Wie kommt mittel- und langfristige Überperformance zustande (Erkenntnisinteresse), und wie kann sie erzeugt werden (normatives Interesse)? Dieses wissenschaftliche und praktische Erkenntnisinteresse ist die Triebfeder für eine breite Zahl an Untersuchungen, die sich mit dem Einfluss verschiedenster Faktoren auf die Leistung eines Unternehmens auseinandersetzen (beispielsweise Peters & Waterman, 1982; Coff, 1999; Ang, 2008; Hull & Rothenberg, 2008; Mackey, 2008).

Gleichzeitig sieht sich das Fachgebiet der Strategischen Managementforschung immer wieder heftiger Kritik ausgesetzt. Bestritten wird insbesondere die Möglichkeit, über Unternehmensgrenzen hinweg übertragbare, also verallgemeinerbare Aussagen zu den Ursachen besonderer Leistungsfähigkeit zu treffen. Der Grund hierfür ist insbesondere in der hohen Komplexität sowie der Spezifik des Einzelfalls zu suchen (March & Sutton, 1997). In einem im deutschen Raum stark beachteten Beitrag halten Alexander Nicolai und Alfred Kieser der Strategischen Managementforschung insbesondere vor, dass es ihr regelmäßig nicht gelingt, einmal gefundene Zusammenhänge in anderen Kontexten zu replizieren, wodurch sie sich in Tautologien flüchten muss, um überhaupt verallgemeinerbare Aussagen zu generieren. Überdies diskutieren die Autoren methodische Schwächen der Operationalisierung¹ der wesentlichen abhängigen Variablen, der Performance der zu untersuchenden Organisationen (Nicolai & Kieser, 2002). Auch verschiedene andere Autoren konstatieren teilweise dramatische Methodenprobleme (vergleiche etwa Short, Ketchen & Palmer 2002 und Boyd, Gove & Hitt, 2005 für einen Überblick).

Diese massiven Kritiken wurden zwar kontrovers, aber durchaus auch zustimmend aufgenommen, die Artikel von March und Sutton sowie Nicolai und Kieser wurden rege zitiert. So weißt Google Scholar, eine Suchmaschine für wissenschaftliche Publikationen, zum 07.01.2010 306 Zitation des Beitrages von March und Sutton und 154 Zitationen der Arbeit von Nicolai und Kieser auf. Die Beiträge gehören damit zwar nicht zum Mainstream, aber

¹ Der Begriff Operationalisierung wird ausführlich in Abschnitt 2.1.4 diskutiert.

doch zum Standardwissen der Managementforschung.² Dennoch lässt Rumelts *problem of persistent profits* die Wissenschaftler trotz aller Schwierigkeiten nicht ruhen und die Strategische Managementforschung erfreut sich weiterhin regen Interesses. Warum aber ist die Strategische Managementforschung nicht in der Lage, reproduzierbare Ergebnisse zu generieren? Ich verfolge in dieser Arbeit die Vermutung, dass sich ein Teil der kritisierten Ergebnisheterogenität aus Methodenproblemen ableiten lässt, die eminent, aber, so meine These, zu überwinden sind.

Die zentralen Schwierigkeiten der Strategischen Managementforschung lassen sich grob in drei stilisierte Schlagworte fassen:

- Die Leistung einer Organisation muss auf eine funktionale Art und Weise operationalisiert werden,
- es müssen relevante Kausalitäten aufgezeigt werden
- und die Ergebnisse müssen über den Einzelfall hinaus generalisierbar sein.

Der Erfassung der zentralen abhängigen Variablen Leistung kommt hierbei eine Schlüsselrolle zu (Richard, Devinney, Yip & Johnson, 2009). Diese kann auf ganz unterschiedliche Art und Weise gemessen werden. So können etwa die Entwicklung des Aktienkurses über einen bestimmten Zeitraum, Rentabilitäten oder die Entwicklung des Marktanteils in einem bestimmten Produkt-/Markt-Segment als Messgröße dienen oder es können Experten befragt oder Umsatzgrößen zu Grunde gelegt werden. Das zentrale Problem hierbei ist, dass die Ergebnisse einer Untersuchung, die die Leistung einer Organisation mit unabhängigen Variablen in Zusammenhang bringt, je nach Operationalisierung der Leistung ganz unterschiedlich ausfallen können. So untersuchen beispielsweise Wiggins & Ruefli (2002) den Zusammenhang zwischen Leistung und Marktanteil und nutzen hierbei verschiedene Operationalisierungen von Leistung. Zum einen wurde der Return on Assets als Grundlage gewählt, zum anderen Tobins q , das Verhältnis des Marktwertes³ eines Unternehmens zu den mit aktuellen Preisen bewerteten Wiederbeschaffungskosten seiner Assets (angenähert durch den Market-to-book-value). In ersterem Fall lag also eine Größe aus dem Bereich Rechnungslegung zu Grunde, in letzterem war insbesondere die Börsenkapitalisierung von Bedeutung. Ein Zusammenhang zwischen Marktanteil und Leistung konnte nun jedoch nur im Falle von Tobins q nachgewiesen werden, ein Zusammenhang zwischen Marktanteil und Return on Assets war nicht festzustellen. Hätten die Autoren lediglich *eine* Leistungsgröße betrachtet, wie regelmäßig üblich, so wären die Ergebnisse vollkommen unterschiedlich ausgefallen: Im Falle einer marktbasier-ten Leistungsgröße wäre ein Zusammenhang zwischen Marktanteil und Performance bestätigt worden, im Falle einer Nutzung des Return on Assets wäre die Hypothese verworfen worden. Hätten also zwei Autoren unabhängig voneinander dieselbe Studie mit unterschiedlichen Leistungsoperationalisierungen durchgeführt wäre exakt der von Nicolai und Kieser kritisierte Fall eingetreten: Die Ergebnisse der ersten Studie wären in der zweiten nicht reproduzierbar gewesen, trotz eines identischen Samples.

² Zum Vergleich: Die Zitationsraten der zusammen mit March & Sutton, 1997 in der gleichen Ausgabe der *Organization Science* (1997, Iss. 6) erschienenen Beiträge liegen zwischen 16 und 186.

³ Wird im Folgenden von Marktwerten, marktbasierter Kennzahlen und ähnlichem gesprochen, ist damit, wenn nicht anders angegeben, der Markt zum Handel mit Anteilsscheinen, also die Börse, gemeint. Marktbasierter Kennzahlen sind also solche, die mit dem Börsenwert eines Unternehmens in Zusammenhang stehen.

Dies wiederum hätte naturgemäß erhebliche Auswirkungen sowohl für das wissenschaftliche Erkenntnisinteresse der Erklärung von Performance-Unterschieden (Einfluss des Marktanteils als Erklärungsgröße von Leistungsunterschiede bestätigt oder verworfen) als auch für die hieraus abgeleiteten normativen Aussagen für die Praxis (Ausweitung des Marktanteils als Strategieempfehlung für Manager). Das gleiche Problem (Market-to-Book Value und ROA) tritt beispielsweise auch bei Marcel (2009) auf – in der Untersuchung zur Zusammensetzung von Top-Management Teams und Performance ändert sich teilweise, im Fall des Diversifikationsgrades, sogar das Vorzeichen der erklärenden Variablen, ebenso bei de Carolis (2003); Combs, Crook & Shook (2005) und Singh, Ang & Leong (2003) nennen eine Reihe weiterer Beispiele. Der weitere Gang der Arbeit wird zeigen, dass die hier anekdotenhaft diskutierten Studien keine Einzelfälle sondern Teil eines verbreiteten Musters sind. Die Gründe hierfür aufzuspüren ist wesentliches Erkenntnisziel der hier vorgelegten Untersuchung.

Offenbar hängt also das Ergebnis von Arbeiten, die sich mit dem Zusammenhang von Leistung und verschiedenen abhängigen Variablen befassen, nicht unwesentlich von der Konstruktoperationalisierung ab. Dabei geht es nicht um die wissenschaftliche Qualität einer Leistungsoperationalisierung sondern um ein gegebenenfalls fehlendes aber zumindest zu wenig stark ausgeprägtes interpersonelles Verständnis davon, was jeweils als Leistung eines Unternehmens gemessen wird. Eine „natürliche“ oder dominante Operationalisierung gibt es dabei nicht, die Breite der Möglichkeiten ist erheblich. Neben den oben angesprochenen aktienkurs-, beziehungsweise accountingbasierten Kennzahlen bieten sich etwa auch qualitative Einschätzungen von internen oder externen Experten (Delaney & Huselid, 1996) oder Faktoren wie das schlichte Überleben eines Unternehmens (beispielsweise in Lamberg, Tikkanen, Nokelainen & Suur-Inkeroinen, 2009) als Leistungsindikatoren an. Es gibt also zahlreiche Möglichkeiten der Operationalisierung von Leistung und die Art der Operationalisierung hat einen wesentlichen Einfluss auf die Forschungsergebnisse im Fachgebiet. Dennoch beklagen gerade bei der Erfassung von Leistung zahlreiche Autoren Defizite über die oben genannte Grundsatzkritik an der Strategischen Managementforschung hinaus und zeigen Forschungsbedarf auf (Chakravarthy, 1986; Lebas, 1995; Zimmerman, 2001; Orlitzky, Schmidt & Rynes, 2003; Powell, 2003b; Boyd et al., 2005; Combs et al., 2005; Devinney, Richard, Yip & Johnson, 2005).

Ein weiteres Problem der Strategischen Managementforschung liegt in der Erklärung von Kausalitäten zwischen Leistung und erklärenden Faktoren. Unternehmen sind dynamische Organisationen, die jederzeit verschiedenen Einflussfaktoren ausgesetzt sind. Umgekehrt haben diese Einflussfaktoren und die Handlungen von Mitgliedern der Organisation Auswirkungen auf ihre Leistung, werden aber nicht alle sofort oder zu einem einfach erkennbaren Zeitpunkt leistungswirksam (Nicolai & Kieser, 2002). So mag sich beispielsweise die Neubesetzung des Top-Managements erst eine, zwei oder drei Perioden nach dem Ereignis in einer bestimmten Leistungsoperationalisierung niederschlagen, wenn die Strategieänderungen des neuen Führungsteams zu tatsächlichen Veränderungen des operativen Ergebnisses führen. Andererseits mag auch bereits die Kommunikation der Neubesetzung zu einer Reaktion der Kapitalmärkte und damit zu einer Indikation bei marktbasierter Leistungsoperationalisierung führen. Zwischen Ursache und Wirkung beziehungsweise Beobachtung der erklärenden Variablen und den möglichen Auswirkungen der Beobachtung auf die erklärte kann also Zeit vergehen, ein Zeitversatz vorliegen (vergleiche auch Chang & Xu, 2008; Henkel, 2009; Sharfman & Fernando, 2008; Short, Ketchen, Palmer & Hult, 2007).

Ein mögliches drittes Problem der Strategischen Managementforschung liegt in der Population⁴, aus der eine Stichprobe gezogen wird, und des aus dieser Stichprobe abgeleiteten Gültigkeitsbereichs der gewonnenen Ergebnisse. Die zentrale Idee jedweder quantitativer Forschung ist die Extrapolation von Mustern, die im untersuchten Sempel zu finden sind, auf eine größere Grundgesamtheit, für die das gezogene Sample⁵ repräsentierend stehen kann (Green, Johnson & Adams, 2006). Der Idealfall wäre, die Gesamtheit aller Unternehmen in ihrer gesamten Tiefe zu jeder beliebigen Zeit betrachten zu können (Carroll, 1994). Die Verwendung einer derartig umfassenden Grundgesamtheit ist in der wissenschaftlichen Praxis naturgemäß nicht möglich, eine Begrenzung auf eine Teilpopulation handhabbarer Größe ist notwendig (vergleiche auch Hitt, Boyd & Li, 2004). Ein weiterer Grund für die Eingrenzung der Grundgesamtheit liegt in der Notwendigkeit, die Zahl dem Forschungsgegenstand exogener Einflussfaktoren, etwa Branche, nationale Herkunft der Untersuchungsobjekte oder deren Größe, zu reduzieren. Entsprechend wird vor dem Ziehen der Stichprobe eine Reduzierung der Grundgesamtheit vorgenommen. Diese Reduktion kann expliziert werden oder implizit mit der Verwendung einer bestimmten Leistungsoperationalisierung einhergehen. So kann die Grundgesamtheit beispielsweise auf Unternehmen einer bestimmten Größe in einer bestimmten Branche einer bestimmten Volkswirtschaft zu einer definierten Zeit beschränkt werden oder die Verwendung einer marktbasierter Leistungsoperationalisierung schließt bereits von vorneherein alle nicht börsennotierten Unternehmen aus. Hiermit geht ein Verlust an Verallgemeinerbarkeit einher, der dann problematisch wird, wenn im Zuge der mit der reduzierten Grundgesamtheit durchgeführten Forschung eine Verallgemeinerbarkeit suggeriert wird, die über die eingeschränkte Grundgesamtheit hinausweist. Es ist keineswegs trivial anzunehmen, dass Ergebnisse einer derart abgegrenzten Teilpopulation aussagekräftig für eine andere Teilpopulation sind, wenn diese sich aus Unternehmen anderer Branchen, Volkswirtschaften oder Größenklasse zusammensetzt. Ergebnisse beziehen sich zunächst immer nur und ausschließlich auf die Teilpopulation, aus der die Stichprobe gezogen wurde, Verallgemeinerungen sind begründungspflichtig. Diese Begründung bleibt jedoch häufig aus, beziehungsweise die eingeschränkte Gültigkeit der Ergebnisse einer Studie wird nicht weiter reflektiert.

Die vorliegende Arbeit befasst sich nun mit diesen drei Problemfeldern der Strategischen Managementforschung: Den Auswirkungen bestimmter Leistungsoperationalisierungen, des unklaren zeitlichen Zusammenhangs zwischen Ursache und Wirkung sowie der Populationsabgrenzung und unzulässiger Rückschlüsse von einer begrenzten auf eine größere Population. Untersucht wird, inwieweit sich diese Problemfelder auf die Ergebnisse von Studien in der Strategischen Managementforschung auswirken. Zu überprüfen ist die Vermutung, dass sich die vielfach kritisch angemerkte Unfähigkeit des Fachgebietes, robuste Kausalitäten zu etablieren, mit diesen Problemfeldern erklären lässt.

⁴ Eine Grundgesamtheit oder Population ist „die Menge der Objekte, für die eine Aussage beabsichtigt ist“ (Schnell, 1991).

⁵ Das Sample beziehungsweise die Stichprobe ist ein selektierter Teil einer Grundgesamtheit, der Rückschlüsse auf diese Grundgesamtheit ermöglichen soll (Hartung, Elpelt & Klösener, 2009).

1.2 Forschungsfragen, zentrales Erkenntnisinteresse, Aufbau und Einordnung

Wie dargestellt ist die Analyse der Einflussfaktoren, die die Leistung von Organisationen erklären können, mit erheblichen Schwierigkeiten verbunden. Im Zentrum dieser Probleme steht die Operationalisierung von Leistung selbst. Verschiedene Autoren untersuchten in der jüngeren Vergangenheit breite Spektren wissenschaftlicher Studien, die sich mit dem Zusammenhang zwischen Leistung und einer oder einer Gruppe von unabhängigen Variablen befassen. So prüfen beispielsweise Orlitzky et al. (2003) in einer umfangreichen Meta-Studie den Zusammenhang zwischen der sozialen und finanziellen Performance von Unternehmen, Combs, Liu, Hall & Ketchen (2006) befassen sich mit dem Zusammenhang zwischen High-Performance Worksystems und Performance und Crook et al. (2008) analysieren die letzten 16 Jahre der Forschung zum Resource-based view. Ein verbindendes Element dieser Meta-Studien ist, dass die Ergebnisse der untersuchten Primärstudien äußerst heterogen ausfallen und wesentlich von der Operationalisierung der abhängigen Variablen Leistung abhängen. Die Implikation: Eine ungeeignete Operationalisierung der abhängigen Variablen kann die Ergebnisse einer Studie nicht nur verzerren, sondern sie im schlimmsten Fall in ihr Gegenteil verkehren, Zusammenhänge suggerieren die nicht existieren oder Kausalitäten, die mit einer anderen Operationalisierung sichtbar geworden wären, verbergen.⁶ Zumindest schmälert dieses Problem wesentlich die Aussagekraft der jeweiligen Studien. Auch eine Populationsselektion, die auf der Begrenzung einer größeren Grundgesamtheit auf ein kleineres Teilsample basiert, kann diese Aussagekraft massiv begrenzen. Schlussendlich arbeiten zahlreiche Studien mit Ein-Perioden-Betrachtungen, es wird also implizit davon ausgegangen, dass die Beeinflussung der abhängigen durch die unabhängigen Variablen in einer Periode wirksam wird. Zahlreiche Autoren äußern Unbehagen ob dieser Annahme (Bowen & Wiersema, 1999; Chang & Xu, 2008; Henkel, 2009; Hmieleski & Baron, 2009 und andere mehr), die ebenfalls eine wesentliche Fehlerquelle und damit eine Erklärung für die beobachtete Heterogenität in den Ergebnissen in der Strategischen Managementforschung darstellen kann.

Zwar wird in verschiedenen Arbeiten auf diese Probleme hingewiesen und Forschungsbedarf aufgezeigt, jedoch liegt derzeit keine Studie vor, die die Probleme Performanceoperationalisierung, Populationsselektion und Zeitversatz in der Strategischen Managementforschung systematisch erarbeitet und insbesondere auf tatsächliche Auswirkungen für Ergebnisse und normative Aussagen des Fachgebietes hin überprüft. Hier setzt das Forschungsvorhaben dieser Arbeit an. Folgende Forschungsfragen und -ziele stehen daher im Fokus:

1. Zunächst ist zu klären, inwiefern sich die Begrenzung der Gesamtpopulation auf homogenere Teilpopulationen auf die Ergebnisse wissenschaftlicher Studien auswirkt und inwieweit die mit der Reduktion auf Teilpopulationen einhergehenden Begrenzungen der Verallgemeinerbarkeit in der Forschungspraxis erkannt und reflektiert werden.
2. Es gilt, einen Überblick über derzeit in der als qualitativ hochwertig anzusehenden Forschung⁷ verwendeten Leistungsoperationalisierungen zu geben.

⁶ Welche Operationalisierung des Leistungskonstrukts geeignet ist, erschließt sich zum einen aus dem Erklärungszusammenhang der jeweiligen Analyse, zum anderen bleibt ein Bewertungsspielraum des durchführenden Wissenschaftlers hinsichtlich seiner Vorstellung einer erstrebenswerten Zieldefinition. Hierauf wird in der Folge detaillierter einzugehen sein.

⁷ Vergleiche hierzu Kapitel 4.

3. Es ist zu klären, wie stark sich unterschiedliche Leistungsoperationalisierungen tatsächlich auf das Ergebnis von wissenschaftlichen Studien auswirken.
4. Ein konzeptionelles Erklärungsraaster, das die Natur des Zusammenhangs zwischen einer bestimmten Leistungsoperationalisierung und einer systematischen Beeinflussung von Studienergebnissen erklären kann, ist zu entwickeln.
5. Zu überprüfen ist im Weiteren, ob das Problem der Time-Lags, der unklaren zeitlichen Zusammenhänge zwischen erklärenden und abhängigen Variablen, die Ermittlung von Kausalaussagen systematisch verzerrt.
6. Schlussendlich gilt es der Frage nachzugehen, welche Empfehlungen sich für die wissenschaftliche Praxis im Fachgebiet der Strategischen Managementforschung ableiten lassen.

Zusammenfassend geht die Arbeit also spezifisch für das Feld der Strategischen Managementforschung der gelegentlich allgemein für die Sozialwissenschaften geäußerten Vermutung nach, dass sich die Schwierigkeit der Aufdeckung kausaler Zusammenhänge und die hieraus resultierende Heterogenität der Ergebnisse empirischer Arbeiten im Fachgebiet ganz wesentlich auf die Methodik der jeweiligen Untersuchungen zurückführen lässt (vergleiche Hunter & Schmidt, 2007).

Die Arbeit gliedert sich in sechs Kapitel. Nachdem in der Einleitung der Forschungsbedarf aufgezeigt wurde und hier nun der Gang der Arbeit vorgestellt wird, gilt es im zweiten Kapitel das Problem des empirischen Erkenntnisgewinns wissenschaftstheoretisch zu beleuchten und dessen Möglichkeiten und Grenzen aufzuzeigen. Der heterogene Begriff der „Leistung“, beziehungsweise „Performance“ im englischen Sprachgebrauch, muss diskutiert und definiert werden. Zudem erfolgt eine Abgrenzung der Leistungsoperationalisierung in der Strategischen Managementforschung gegenüber dem Business Performance Measurement, der Leistungsmessung zum Zwecke der Unternehmenssteuerung, sowie eine knappe Vorstellung des in dieser Arbeit interessierenden Fachgebiets der Strategischen Managementforschung und des derzeit für das Fachgebiet relevanten epistemologischen Zugangs. Das zweite Kapitel liefert somit den Begründungszusammenhang der Arbeit.

Das dritte Kapitel der Arbeit erarbeitet den theoretischen Hintergrund der in der Arbeit diskutierten Probleme der Strategischen Managementforschung, Ansätze zur Erklärung eines möglichen systematischen Einflusses von Leistungsoperationalisierungen auf Forschungsergebnisse in der Strategischen Managementforschung werden entwickelt. Basis für die Erklärung des Einflusses der Leistungsoperationalisierung ist insbesondere die Appropriationstheorie, die, aus dem Bereich der Stakeholder-Forschung stammend, ich theoriebildend auf die Problemlage der Leistungsoperationalisierung anwende. Das Problem der Verallgemeinerung von Ergebnissen aus Teilpopulationen auf eine größere Population wird ebenso aus einem theoriegeleiteten Verständnis analysiert wie die methodische Schwierigkeit des unklaren zeitlichen Bezugs zwischen Ursache und Wirkung.

Anschließend folgt der empirische Teil der Arbeit, der die Kapitel 4 und 5 umfasst. Im Rahmen der vorliegenden Dissertationsschrift führe ich zwei separate empirische Studien durch. In Kapitel 4 wird im Rahmen einer Literaturreview, die jeweils drei Jahrgänge zweier führender Journals im Bereich der Managementforschung auswertet, die Problemlage deskriptiv beleuchtet. Auf dieser Basis aufbauend folgt in Kapitel 5 eine umfangreiche Metaanalyse, die

anhand einer quantitativen Auswertung vorliegender Primärstudien zum Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen die in Kapitel 3 entwickelten Hypothesen testet.

Im abschließenden 6. Kapitel diskutiere ich die gewonnenen Ergebnisse und ordne sie in den theoretisch-konzeptionellen Bezugsrahmen ein. Schließlich werden Empfehlungen zum Umgang mit Methodenproblemen in der Strategischen Managementforschung im Allgemeinen und der Operationalisierung von Leistung im Besonderen abgeleitet. Forschungsdefizite und mögliche Pfade für Folgeforschungen werden aufgezeigt.

Methodisch und thematisch ähnelt die Arbeit damit den Untersuchungen von Schilling, (2009), Capon, Farley & Hoenig (1990), Short et al. (2002) und Richard et al. (2009). Schilling untersucht die Eignung verschiedener Datenbanken zur Analyse von Allianzen und mögliche Auswirkungen der Verwendung bestimmter Datenbanken auf das Ergebnis von Untersuchungen und beginnt ebenfalls mit einer deskriptiven, gefolgt von einer metaanalytischen Studie. Ähnlich Capon und Kollegen, die Beziehungen zwischen nichtfinanziellen Faktoren und finanzieller Performance in einer Meta-Analyse studieren. Short und Kollegen untersuchen den Einfluss der Populationsselektion auf die Ergebnisse von Studien in der Strategischen Managementforschung und nutzen hierzu ebenfalls eine Literaturreview und eine weitergehende quantitative Analyse. Inhaltlich folgt die Arbeit dem Forschungsstrom der Methodenentwicklung in der Strategischen Managementforschung, der ausführlich von Ketchen, Boyd & Bergh (2008) vorgestellt wird. Richard et al. (2009) führen eine der in Kapitel 4 präsentierten Literaturreview ähnliche Analyse durch.

Die vorgelegte Arbeit trägt zum Wissensfortschritt im Fachgebiet bei, in dem

- (1) deskriptive Ansätze zur Untersuchung der Leistungsoperationalisierung in der Strategischen Managementforschung von Richard et al. (2009) oder March & Sutton (1997) mit metaanalytischen Untersuchungen ähnlich Short et al. (2002) vereint werden,
- (2) ein theoretisches Gerüst zur Analyse von Methodenproblemen in der Strategischen Managementforschung und insbesondere zum Einfluss der Leistungsoperationalisierung auf Studienergebnisse entwickelt wird und
- (3) aus theoretischen Überlegungen zu Methodenproblemen testbare Hypothesen entwickelt und diese einer Überprüfung auf empirische Relevanz unterzogen werden.

Eine vergleichbare Studie ist meines Wissens in der Strategischen Managementforschung so bisher nicht durchgeführt worden, hieraus begründet sich die Notwendigkeit und wissenschaftliche Relevanz der vorgelegten Arbeit.

2. Grundlagen

2.1 Die Rolle von Theorie, Empirie und Hypothese in der Wissensgenerierung

2.1.1 Induktion, Deduktion und das Problem der Verallgemeinerbarkeit

Zu Beginn einer Arbeit, die sich mit Methodenproblemen befasst, muss notwendigerweise die Art der Wissensgenerierung Beachtung finden, vor deren Hintergrund die zu diskutierenden methodischen Schwierigkeiten auftreten.⁸

Grundsätzlich lässt sich der Erkenntnisauftrag der Wirtschaftswissenschaften in einen erfahrungs- und einen handlungswissenschaftlichen Bereich teilen. Im Rahmen der Erfahrungswissenschaft gilt es, die komplexen Zusammenhänge des realen wirtschaftlichen und sozialen Lebens in mehr oder weniger einfachen Erklärungsansätzen beherrschbar zu machen und kausale Zusammenhänge aufzuzeigen. Man bedient sich hierbei eines axiomatischen Systems, das heißt es werden aus möglichst wenigen Grundannahmen (Axiomen⁹) möglichst einfache Sätze (Aussagen) über Zusammenhänge gebildet. Diese Sätze, beziehungsweise aus diesen abgeleitete Hypothesen, werden sodann durch Konfrontation mit der Empirie auf ihre Eignung zur Erklärung der komplexen Realität hin überprüft. Zu einem gemeinsamen Denkkonstrukt zusammengefasste Sätze, die auf einem konsistenten Axiomensystem aufbauen, bilden Theorien (vergleiche Maurer, 2004, 11f). Auf der Seite der Handlungswissenschaften gilt es, die gefundenen, beziehungsweise in der Empirie bestätigten Sätze in geeigneter Weise zur Erreichung von Zielen einzusetzen, beziehungsweise aus diesen Sätzen gewonnene Erkenntnisse als Wege zu Zielen anzubieten. Gewonnen werden können Theorien grundsätzlich sowohl induktiv als auch deduktiv, also sowohl dadurch, dass aus Beobachtungen der Realität, „besonderen Sätzen“, auf diesen zu Grunde liegende Gesetzmäßigkeiten, „allgemeine Sätze“, geschlussfolgert wird¹⁰, als auch durch theoretische Annahmen über diese Gesetzmäßigkeiten, die dann wiederum in der Realität zu überprüfen sind (Leibniz, 1704/1961), also Schlüssen von allgemeinen Sätzen auf besondere. Fraglich ist hierbei, ob rein induktiver Wissensgewinn grundsätzlich möglich ist, interagiert der Forschende doch unweigerlich mit dem zu untersuchenden Objekt. Die Persönlichkeit des Wissenschaftlers ist von seinem Erkenntnisstreben nicht trennbar (Miller, 2005). Mit diesem Befund einher geht die Erkenntnis, dass

⁸ Hier muss vorausgeschickt werden, dass ich in der vorliegenden Arbeit einen kritisch-rationalistischen Ansatz verfolge. Gelegentliche Blickpunktwechsel in Richtung konstruktivistischer Sichtweisen begründe ich an den entsprechenden Stellen und kennzeichne sie entsprechend. Wie die weiteren Ausführungen zeigen werden ist diese Sichtweise in der Strategischen Managementforschung in ihrer Reinform nicht uneingeschränkt anwendbar und wird entsprechend auch im Folgenden kritisch diskutiert. Trotz der offensichtlichen Probleme kritisch-rationalistischer Sichtweisen wähle ich diesen Blickwinkel, da er sowohl was die Forschungspraxis als auch was die denkgeschichtliche Dimension angeht die stärkste Akzeptanz im derzeitigen Diskurs des Fachgebietes findet (Durand & Vaara, 2009). Eine umfassende Diskussion verschiedener epistemologischer Zugänge zur Strategischen Managementforschung liegt jenseits des Fokus dieser Arbeit, der interessierte Leser sei auf Durand & Vaara (2009) verwiesen.

⁹ Der in der aktuellen wissenschaftstheoretischen Diskussion grundgelegte Begriff des Axioms ist nicht mit dem klassischen Axiombegriff zu verwechseln, wie er etwa von Leibniz benutzt wird. Ein Axiom ist im hier zu Grunde gelegten Verständnis (Wolf, 2008) jede nicht abgeleitete Aussage einer Theorie, während im klassischen Verständnis ein Axiom definiert war als ein Satz, der „aus sich selbst heraus evident“ ist (Leibniz, 1704/1961, 352).

¹⁰ Zitation aus Popper, 1989, 3.

das Forschungsideal eines von der Person des Forschers vollständig unabhängigen Ergebnisses grundsätzlich nicht vollumfänglich erreicht werden kann. Entsprechend kann es keine objektiven Beobachtungen geben und es ist nicht möglich, von Beobachtungen auf Theorien zu schlussfolgern, ohne dass zumindest Ansätze dieser Theorien in Form von Ideen bereits ex ante vorhanden sind und damit in die Beobachtung eingeflossen wären. Der Grund hierfür ist, dass sich der Wissenschaftler in einem sozialen und kognitiven Rahmen bewegt, der ihm den Reflexionsraum für seine Erkenntnisse und das Vorwissen, sei es kodifiziert oder tacit, zur Verfügung stellt (Kuhn, 2007; Foucault, 1970/2003)¹¹. Durch die Interaktion des Wissenschaftlers, bzw. seiner Theorien, mit dem zu untersuchenden Objekt ist es möglich, dass sich unterschiedliche Denkstile, oder Paradigmen in der Kuhn'schen Terminologie, ihre eigenen Wirklichkeiten konstruieren (Schumann, 2006; Lehmann-Waffenschmidt, 2006). Dies bedeutet nicht zwingend den Rückgriff auf ein radikalkonstruktivistisches Weltbild und damit eine Abkehr von der erklärenden Wissenschaft im Popperschen Sinne, allerdings muss anerkannt werden, dass in der Interpretations- und Analysearbeit die Werturteile des Forschers und des Paradigmas oder Denkstils, in dem er sich bewegt, zwangsläufig auf die Ergebnisse seines Erkenntnisstrebens einwirken. Wechsel zwischen Paradigmen sind mit erheblichem Aufwand und Kosten verbunden (Kuhn, 2007). Wissensgenerierung ist demnach also immer deduktiv, fraglich ist nur, inwieweit die zu Grunde liegenden Theorien expliziert werden. Eine umso größere Rolle kommt der Explikation von Theorie als Ausgangspunkt des Wissensschöpfungsprozesses zu. Durch diese Explikation kann sie auch in anderen Denkkontexten überprüft werden, ihr empirischer Gehalt, ihr „Bewährungsgrad“ steigt, mit der „Strenge der Prüfung, der [sie] (bis zu einem bestimmten Zeitpunkt unterworfen wurde“ (beide Zitate aus Popper, 1989, S. 211).¹²

¹¹ Diese Argumentation wiederum deckt sich mit Poppers These vom Fortschritt in „induktiver Richtung“, bzw. der „Quasiinduktion“ (Popper, 1989, 221).

¹² Neben dem hier vorgestellten Argument wird häufig auch das Argument des unendlichen Regresses als Begründung für die Unmöglichkeit induktiver Wissensgenerierung angeführt (vergleiche beispielsweise Popper, 1989). Popper argumentiert, dass, um aus Induktion auf Theorien zu schließen, Verfahren („Induktionsprinzipien“, S. 5) notwendig sind, die vorgeben, wie solche Schlüsse aussehen müssen. Diese Verfahren wiederum müssen einer eigenen induktiven Logik folgen, die ebenfalls zu begründen ist. Diese Logik wird, wenn sie induktiv gewonnen wird, ebenfalls wieder auf andere induktive Logiken zurückzuführen sein und so weiter. Es ergibt sich also ein endloser Regress. Die Schlussfolgerung aus der Beobachtung auf die Theorie ist daher logisch nie belastbar möglich, da die Prinzipien, nach denen dieser Schluss erfolgen müsste, diesem unendlichen Regress unterworfen und damit nie festlegbar oder greifbar sind, es sei denn durch die dogmatische Festlegung einer letztlich nicht begründbaren Grenze des Regresses. Diese Überlegung ist aus meiner Sicht nicht vollständig überzeugend, da Popper der sich aufdrängenden Gegenfrage, ob denn deduktive Methodik ohne solche Regresse auskommen kann, nur entgeht, indem er den Prozess der Ideengenerierung, also die Entwicklung der durch Deduktion zu prüfenden Sätze, definitorisch als zur Psychologie und nicht zur Wissenschaftstheorie gehörig ausschließt (Popper, 1989). Aus diesem Grunde arbeite ich im Folgenden argumentativ nicht mit Popper sondern mit Foucault und Kuhn. Das zweite formallogische Argument Poppers gegen induktives Vorgehen, seine Kritik am „Psychologismus“, soll hier aufgrund seiner aus meiner Sicht eher geringen Relevanz für die hier untersuchte Fragestellung aus Fokussierungsgründen nicht weiter verfolgt werden und kann bei Popper, 1989, 60f, nachgelesen werden. Ein relevantes Argument gegen rein induktives Vorgehen in empirischen Wissenschaften findet sich bei Hunter & Schmidt (2007), die davon ausgehen, dass sich ein wesentlicher Teil der Heterogenität von Forschungsergebnissen durch statistische Artefakte erklären lässt. Deduktives Vorgehen ermöglicht den Test von Theorien und damit die Beurteilung, ob eine Beziehung tatsächlich von Relevanz ist oder lediglich als statistisches Artefakt in die Daten Eingang gefunden hat. Bei induktivem Vorgehen ist eine solche Unterscheidung nicht möglich, sinnvolle Ergebnisse sind empirisch so kaum zu erlangen.

Der durchaus berechtigten gegenläufigen Argumentation, dass die Bildung synthetischer, also sich nicht aus sich selbst heraus erklärender, Sätze und Theorien ohne Rückgriff auf induktive Erfahrungswerte nicht möglich ist, setzt Kant seine Transzendentalphilosophie entgegen (Kant & Eisler, 1930/2009). Auch hier ist die Interaktion von Objekt und Subjekt, von Untersuchungsgegenstand und Wissenschaftler, entscheidend. Wichtig für die Gewinnung von Erkenntnis ist demnach aber nicht die Betrachtung des Objektes, also des Untersuchungsgegenstandes, die, da wir die Außenwelt nicht als gegeben voraussetzen können, keine unmittelbaren Erkenntnisse erbringen kann, sondern vielmehr die „Selbstschöpfung der Ideen zu einem vollständigen System der Gegenstände der reinen Vernunft“ (zitiert nach Kant & Eisler, 1930/2009, Stichwort „Transzendentalphilosophie“). Als Theorie der Metaphysik entzieht sich Kants Gedankengebäude wiederum der empirischen Überprüfung. Dies ist konzeptionell etwas unbefriedigend, da ja Popper die induktiven Vorgehensweisen explizit mit Werkzeugen der Logik kritisiert und mit diesen argumentiert (vergleiche Fußnote 12). Dennoch kann es als wissenschaftlicher Konsens gelten, dass es tatsächlich möglich und üblich ist, deduktiv zu arbeiten, während rein induktive Verfahren in der hochwertigen Forschung kaum vorkommen.¹³ Wissenschaftsgebiete, die sich dem induktiven Zugang entziehen (etwa das gesamte Themenfeld der theoretischen Physik) und dennoch auf breiten Theoriegebäuden ruhen, bestärken die Kant'schen Thesen (vergleiche auch hierzu Popper, 1989).

Ansammlungen von Annahmen, Sätzen und Ideen müssen einigen Grundanforderungen genügen, um als Theorien zu gelten: Zunächst müssen die Axiome, die die Grundlage der Theorie bilden sollen, untereinander widerspruchsfrei und dürfen zueinander nicht redundant sein (Konsistenz- und Unabhängigkeitspostulat). Theorien müssen so weit wie möglich vollständig sein, also die wesentlichen Einflussfaktoren und Kausalitäten abdecken (Vollständigkeitspostulat) und für einen weiten Kreis von Anwendungsfeldern Gültigkeit besitzen (Allgemeingültigkeitspostulat). Theorien müssen die Ableitung klarer Wenn-Dann-Aussagen zulassen (Bestimmtheitspostulat) und müssen empirisch gehaltvoll sein, das heißt eine möglichst große Zahl an Fällen ausschließen. Damit zusammenhängend müssen sie falsifizierbar, das heißt widerlegbar, sein (Popper, 1989) und Gesetzescharakter haben, ein Auslöser soll also zwingend zu einem bestimmten Ergebnis führen (Wolf, 2008, 14-18).¹⁴

Fraglich ist, inwieweit Theorien, die diesen Anforderungen nicht vollständig gerecht werden, abzulehnen sind. Im Sinne eines erklärenden Wissenschaftsverständnisses sind Theorien so lange weiterzuentwickeln, bis sie diesen Grundforderungen entsprechen. Dies geschieht durch den Prozess der Falsifikation (Popper, 1989, siehe unten). Im verstehenden oder hermeneutischen Wissenschaftsverständnis wird diese Forderung nach Ablehnung einer falsifizierten Theorie relativiert und als für die Sozialwissenschaften nicht absolut einhaltbar gewertet. Insbesondere die Zeit- und Umstandsbezogenheit der Geistes- und Sozialwissenschaft-

¹³ Dies ist zumindest im hier zu untersuchenden Fachgebiet der Strategischen Managementforschung der Fall, wie die Literaturreview, die in Kapitel 4 präsentiert wird, zeigt. Die Arbeitsweise ist durchgängig deduktiv im popperschen Sinne: Aus einem Theoriegebäude werden Hypothesen abgeleitet, die sodann empirisch überprüft werden. Dennoch hat die induktive Methode auch in der aktuellen Forschung mit dem Konzept der Grounded Theory (Glaser & Strauss, 1977) zumindest nominell noch relevante Vertreter, auch wenn diese Methode zumindest in dem Ausschnitt des Forschungsgebietes, der im Rahmen der Literaturreview betrachtet wurde, keine Rolle spielt.

¹⁴ Vergleiche zur Rolle von Axiomen in der Theoriebildung beispielsweise auch die Diskussion von Goshal (2005) des Einflusses des radikalindividualistischen Menschenbildes auf die Neoinstitutionalistischen Organisationstheorien.

ten wird hierbei als wesentlicher Faktor eingeführt. Demnach ist in den Sozialwissenschaften nie eine vollständig replizierbare Kausalität zu erwarten, vielmehr sind Zusammenhänge grundsätzlich historisch und kontextgebunden (Gadamer, 1990, hierauf wird in Abschnitt 4.2 noch genauer einzugehen sein). Gadamer betont die Unterschiede zwischen Geistes- und Naturwissenschaften (die Popper (1989) und insbesondere Maurer (2004) bestreiten) und lehnt eine Vereinheitlichung der Erkenntnistheorien aller Erkenntniswissenschaften, ein „Ideal einer Naturwissenschaft der Gesellschaft“ (S. 10), wie sie dem Gedankengebäude Poppers zu Grunde liegt, ab.¹⁵ Damit wird die grundsätzliche Möglichkeit, dass eine Theorie in den Geisteswissenschaften (stabilen) Gesetzescharakter aufweisen könnte, verworfen, ebenso das Allgemeingültigkeitspostulat. Wolf (2008) bezeichnet die Forderungen des erklärenden Wissenschaftsverständnisses diesem Gedanken folgend für die Sozialwissenschaften als „Fiktion“ (S. 22). Gleichzeitig verweist er jedoch auf der anderen Seite auch auf die faktische Irrelevanz, die wissenschaftlichen Aussagen aus dem Kontext des hermeneutischen Verständnisses heraus zu Eigen ist, da dieses Verallgemeinerbarkeit ablehnt. Damit müssen Aussagen auf dem Niveau eines „story-tellings“ (S. 23) bleiben, bilden also immer nur die Basis zur Betrachtung von Einzelfällen. Theorien im oben definierten Sinne können so nicht entstehen.

Gadamers Kritik kann durchaus als Vorläufer der Grundsatzkritik an der Strategischen Managementforschung verstanden werden, wie sie von March & Sutton (1997) vorgebracht wurde und in der Einleitung sowie in Abschnitt 2.3 diskutiert wird. Es stellt sich die Frage, ob in einer hochkomplexen, sich ständig verändernden Welt realistisch übertragbare, das heißt in unterschiedlichen Settings und Zeiten gültige Kausalitäten vorstellbar sind. Dieser Dissens kann an dieser Stelle nicht aufgelöst werden und wird im Rahmen von Abschnitt 2.3 noch weiter zu diskutieren sein. Die grundsätzliche Existenz einer theoretisch fundierten Sozialforschung spricht dafür, dass in weiten Teilen der wissenschaftlichen Gemeinschaft zumindest ein sozialer Konsens im Kuhn'schen Sinne darüber besteht, dass die Suche nach Theorien, die dem Allgemeingültigkeitsprinzip zumindest in gewissen Grenzen genügen, nicht hoffnungslos ist.

Festzuhalten bleibt, dass Theorien einigen Grundanforderungen genügen müssen, um als solche im Rahmen der Wissensgenerierung relevant zu sein. Wichtig ist hierbei insbesondere der Anspruch (wenn auch nicht zwingend in absoluten Maßstäben) einer Übertragbarkeit, das heißt ein System an Sätzen lässt sich nur dann als Theorie bezeichnen, wenn es Aussagen generiert, die nicht nur für eine bestimmte Situation sondern für ein breites Spektrum an Situationen von Interesse sind. Der Wert einer Theorie steigt mit zunehmendem empirischem Gehalt, also mit der Zahl der durch die zu prüfenden Sätze ausgeschlossenen Fälle, und zunehmender Übertragbarkeit, also mit der Weite an Kontexten, in denen diese Sätze überprüft werden. Das Problem der Übertragbarkeit von Ergebnissen über Teilpopulationen hinweg wird im weiteren Verlauf der Arbeit als möglicher Grund für Heterogenitäten in den Ergebnissen im Forschungsgebiet der Strategischen Managementforschung, dem Entdeckungszusammenhang dieser Arbeit, noch relevant werden.

¹⁵ Nicht jedoch eine Vereinheitlichung der Methoden. Diese müssen nach Gadamer, 1990, 13, in allen Wissenschaften die selben sein und sind nur in den Naturwissenschaften am prägnantesten ausgeprägt.

2.1.2 Wissensfortschritt durch Falsifikation

Wichtig ist nun insbesondere, wie eine Theorie auf Relevanz überprüft werden kann. Ein solcher Theorietest ist nötig, da sozialwissenschaftliche Theorien Erklärungen der Realität anstreben. Eine solche Erklärung muss sich an eben der Realität messen lassen, die sie zu erklären sucht (Maurer, 2004, 17), sie muss, in den Worten von Popper, 1989, 15, „an der Erfahrung scheitern können“. Aus den Sätzen, die eine Theorie bilden, leiten sich Hypothesen ab, Aussagen darüber, wie eine Beziehung ausfallen müsste, wenn die Theorie wahr wäre. Diese Hypothesen können nun an der Realität durch Konfrontation mit Beobachtungen überprüft werden. Wichtig ist, dass auch die Bestätigung einer Hypothese die Theorie nicht als wahr kennzeichnen kann (Leibniz, 1704/1961, 565; Popper, 1989, 3). Der Grund hierfür ist, dass nie die Existenz einer falsifizierenden Beobachtung (Poppers viel zitiertes „schwarzer Schwan“, vergleiche Popper, 1989, 4-8) ausgeschlossen werden kann. Zudem kann es immer eine Theorie geben, die die beobachtete Realität ebenso gut erklärt wie die überprüfte, zudem jedoch auch noch einen weiteren Teil der Realität erklärt, der in die Ursprungstheorie nicht einbezogen war. Diese Erkenntnis der Unmöglichkeit der Verifikation von Theorien ist als Hume'sches Induktionsproblem in die Wissenschaftsgeschichte eingegangen (Maurer, 2004; Popper, 1973). Aus dem Hume'schen Induktionsproblem folgt, dass auch eine Konfrontation mit der Empirie nicht endgültig über Wahrheit und Falschheit einer Theorie entscheiden kann. Damit tritt bei derartigen Theorietests eine „logische Asymmetrie“ (Maurer, 2004, 20) auf: Wird eine empirisch überprüfbare Hypothese durch Konfrontation mit der Empirie bestätigt, lässt sich hieraus nicht schließen, dass die zu Grunde liegende Theorie beziehungsweise ihr Axiomensystem wahr ist, umgekehrt wirkt die Falsifikation einer Hypothese jedoch sehr wohl auf die Theorie beziehungsweise das zu Grunde liegende Axiomensystem zurück (Popper, 1989, 211). Da eine Hypothese logisch aus den Axiomen der Theorie abgeleitet ist, muss mindestens eines der Axiome falsch sein, wenn die Hypothese verworfen wird.

Wie kann aber nun Wissenszuwachs stattfinden, wenn Theorien nie verifiziert werden können und Wissenschaft wesentlich auf der Entwicklung von Theorien beruht? Nach Popper findet Wissenszuwachs nicht auf dem Wege der Bestätigung von Theorien sondern durch ihre Falsifikation statt: Wird eine Theorie falsifiziert, tritt eine neue an ihre Stelle, die erklären kann, was auch die falsifizierte Theorie erklären konnte, zuzüglich weiterer Phänomene, an denen die alte Theorie scheiterte (Popper, 1973). Die neue Theorie kann dabei vollständig unabhängig von der alten Theorie sein oder diese weiterentwickeln (Maurer, 2004), wobei Popper davon ausgeht, dass letzteres der Regelfall ist (Popper, 1989, 199). Ein Problem tritt insbesondere dann auf, wenn nach Falsifikation der Theorie A keine Theorie B verfügbar ist, die eben den oben dargelegten Kriterien entspricht, die Realität also besser erklären kann als die falsifizierte Theorie. Poppers kritischer Rationalismus betont, dass eine Theorie immer dann falsch ist, wenn mindestens eine aus ihr abgeleitete Hypothese falsifiziert ist. Damit wäre in einem solchen Fall Theorie A zu verwerfen und nach einer neuen Theorie zu suchen. Die alte Theorie könnte bestenfalls noch als „Grenzfall der neuen Theorie“ (Popper, 1989, 199) weiterexistieren. Diese neue Theorie wird so lange beibehalten, bis sie wiederum falsifiziert wird und eine weitere Theorie entstehen muss, die Systematik setzt sich fort und das Wissen entwickelt sich stetig weiter. Man nähert sich einer als existent angenommenen objektiven Realität immer weiter an („quasiinduktiver“ Wissensfortschritt, Popper, 1989, 221). Wichtig ist hierbei ein Wettstreit der Theorien, aus denen die besser bewährte als näher an der tatsächlich existenten äußeren Wirklichkeit als Sieger hervorgeht.

Fraglich ist, inwieweit die Forderung nach einem Verwerfen falsifizierter Theorien in den Sozialwissenschaften, die, wie Gadamer (1990) ausführte, ja durch Historizität und Ambiguität geprägt sind, funktional sein kann. Die Gegenposition hierzu diskutierte wohl am einflussreichsten Milton Friedman in seinem Beitrag „The Methodology of Positive Economics“ (Friedman, 1966/1994). Eine Theorie wäre diesem instrumentalistischem Gedankengang nach immer dann beizubehalten, wenn sie „valid and meaningful (i.e. not truistic) predictions about phenomena“ generiert (S. 7), also die Realität befriedigend erklären kann. Eine Übereinstimmung mit einer objektiv existenten Realität ist hierbei kein Kriterium, im Gegenteil ist es häufig sogar nötig, unrealistische und damit leicht falsifizierbare Annahmen über die Realität zu treffen, um zu den geforderten validen und relevanten Aussagen über die Realität zu kommen. Ein Beispiel hierfür ist das Modell der vollständigen Konkurrenz, dessen Axiome leicht in Hypothesentests zu falsifizieren sind (etwa die Annahme vollständiger Information aller Marktteilnehmer), das aber dennoch eine brauchbare Erklärung vieler Phänomene liefert und deshalb nach wie vor nicht verworfen ist (das Beispiel ausführlich: Maurer, 2004, 15). Für das hier interessierende Fachgebiet der Strategischen Managementforschung argumentiert insbesondere Powell ähnlich (Powell, 2001; Powell, 2002). Aufgrund von kausaler Ambiguität und vager Begrifflichkeiten und Konstrukte (etwa Leistung, Performance, Ressource, vergleiche den weiteren Gang der Arbeit) sind regelmäßig keine klaren Zusammenhänge zu erwarten. Vielmehr ist damit zu rechnen, dass, wenn eine Beziehung überprüft wird, eine große Zahl der spezifischen Untersuchung exogener Faktoren die fragliche Kausalität beeinflussen; damit wären reproduzierbare Ergebnisse kaum zu erwarten.¹⁶ Hunter & Schmidt (2007) argumentieren im Weiteren, dass Methodenprobleme, so genannte statistische Artefakte, für einen wesentlichen Teil der Heterogenität der Ergebnisse in sozialwissenschaftlichen Studien verantwortlich sind. Diese Methodenprobleme lassen sich allerdings nie völlig vermeiden sondern lediglich eingrenzen. Hierdurch wäre es unsinnig, die Hypothesenfalsifikation als Ablehnungsgrund einer Theorie heranzuziehen, da zunächst unklar ist, ob die Falsifikation nicht durch methodische Schwierigkeiten zu Unrecht erfolgte.¹⁷

Wesentliches Kriterium für das Verwerfen von Theorien wäre demnach nicht die Falsifikation von Hypothesen an einem Einzelfall (Poppers Schwarzer Schwan), sondern die Begutachtung der Nützlichkeit, die eine Theorie für den entsprechenden Fachdiskurs entfalten kann. Wird diese Nützlichkeit nicht mehr gesehen, ist die Theorie zu verwerfen, solange dies nicht der Fall ist, wird sie beibehalten. Mit Powells Argumentation ist auch die aus kritisch-rationalistischer Sicht absurde Beobachtung zu erklären, dass in den Sozialwissenschaften offenkundig inkompatible Theorien zur Erklärung des gleichen Phänomens nebeneinander existieren, wie etwa im Fall des Neoklassischen und des Schumpeterschen Marktmodells (Schumpeter, 1952).¹⁸

Die Praxis der Strategischen Managementforschung bewegt sich zwischen diesen Polen des kritischen Rationalismus auf der einen und des reinen Instrumentalismus auf der anderen

¹⁶ Vergleiche hierzu auch die Diskussion zur Organisationskybernetik als alternativem Zugang zur Strategischen Managementforschung in Abschnitt 2.3.

¹⁷ Vergleiche zu den von Hunter und Schmidt diskutierten Methodenproblemen auch Anhang A.3.2.1.

¹⁸ Beispiele hierfür finden sich auch im hier besonders interessierenden Fachgebiet der Strategischen Managementforschung, etwa in der Kontroverse zwischen Market- und Resource-based view (vergleiche beispielsweise Hult & Ketchen, 2001; Peteraf & Barney, 2003). Vergleiche hierzu auch die Ausführungen weiter unten in diesem Abschnitt.

Seite. Angesichts der Tatsache, dass Arbeiten in der Strategischen Managementforschung regelmäßig zu äußerst heterogenen Ergebnissen kommen (vergleiche etwa die beiden Meta-Studien zum Resource-Based View:¹⁹ Newbert (2007); Crook et al. (2008)) ist klar, dass, bei allem Wettstreit der Theorien, Poppers Forderung nach einem konsequenten Verwerfen von Theorien, die einmal falsifiziert sind, in der Strategischen Managementforschung kaum durchzuhalten sein kann. Insbesondere die aufgrund der hohen Komplexität immer zahlreich vorhandenen unbeobachteten Variablen verhindern grundsätzlich eine so eindeutige Spezifikation von Theorien, dass diese durch replizierbare Hypothesentests vollständig so falsifiziert werden könnten, dass die Gründe der Falsifikation eindeutig den Sätzen der Theorie und nicht der Theorie exogenen Faktoren wie statistischen Störgrößen beziehungsweise Moderatorvariablen zugeschrieben werden könnten. Vielmehr müssen sozialwissenschaftliche Theoriegebäude der kausalen Ambiguität durch eine gewisse Aussagenunschärfe Rechnung tragen (vergleiche hierzu auch Abschnitt 4.2). Entsprechend ist das absolute Falsifizierungsgebot im Sinne Poppers in den Sozialwissenschaften ebenso wenig sinnvoll wie die oben diskutierte Forderung nach dem Gesetzescharakter von Theorien. Gleichzeitig droht der reine Instrumentalismus in Beliebigkeit abzugleiten und eine theoretische Fundierung beziehungsweise deduktive Wissensgewinnung hinfällig werden zu lassen. Die Probleme, die mit einer Abkehr von der deduktiven Perspektive verbunden sind, wurden bereits in Abschnitt 2.1.1 diskutiert. In den Sozialwissenschaften spricht man folgerichtig von der Robustheit einer Theorie in dem Sinne, als dass diese empirisch in der Mehrzahl der Fälle Bestätigung findet und einen Erklärungsbeitrag für eine Problemstellung liefern kann (Hall & Tickle-Degnen, 1994). Zur Annäherung an die als gegeben angenommene objektive Wahrheit bietet sich das Methodenspektrum der Wissenssynthese an (Cooper & Hedges, 1994), das auch die empirische Basis der vorliegenden Arbeit bildet.

Realistisch zu erwarten ist also eine Arbeit mit Theorien, die, obschon in einigen Kontexten und Studien falsifiziert, die von Friedman geforderten validen und relevanten Aussagen liefern können, so lange, bis sie von einer anderen, auch im popperschen Sinne besseren Theorie abgelöst oder, häufiger, ergänzt werden. Dies zeigt sich etwa an der nach wie vor geführten Debatte, ob der seit Mitte der 1990er Jahre in der Strategischen Managementforschung dominierende Resource-based view den insbesondere in den 1980er und frühen 1990er Jahren dominanten Market-based view ersetzt hat oder ersetzen sollte. Beide Theorien versuchen die Überperformance von Unternehmen zu erklären. Der ältere Market-based view erklärt Performanceunterschiede insbesondere mit mehr oder weniger erfolgreichem strategischem Marktverhalten eines Unternehmens, während der Resource-based view Performanceunterschiede auf unterschiedliche wertschöpfungsrelevante Ressourcen innerhalb eines Unternehmens und deren erfolgreicher Nutzbarmachung zurückführt (Porter, 1999; Barney, 1991; Peteraf, 1993). Wiggins & Ruefli (2002) falsifizieren die Theorie des Market-based view und plädieren kritisch-rationalistisch entsprechend dafür, die Kernaussagen die-

¹⁹ Die Theorie des Resource-based view kann als der derzeit dominante Ansatz zur Erklärung nachhaltiger Unterschiede in der Leistung von Organisationen gesehen werden (Powell, 2001; Newbert, 2008; Crook, Ketchen, Combs & Todd, 2008). Aus diesem Grund wird in dieser Arbeit häufig auf diese Theorie Bezug genommen, wenn forschungspraktische Beispiele in der Strategischen Managementforschung Beachtung finden sollen. Dies geschieht primär aus Gründen der Erklärungs Klarheit der verfolgten Argumentationen heraus, die diskutierten Problemstellungen finden sich in ähnlicher, wenn auch weniger prominent diskutierter Form auch in anderen Erklärungsmodellen. Für einen Überblick über Alternativen zum Resource-based view vergleiche beispielsweise Teece, Pisano & Shuen, 1997.

ser Theorie zu verwerfen. Gleichzeitig ist jedoch auch der Resource-based view verschiedentlich falsifiziert worden, auch wenn diese Theorie in Summe ein erheblicher Erklärungsbeitrag zugestanden wird (vergleiche die bereits oben zitierten Arbeiten von Wiggins & Ruefli, 2002; Newbert, 2007 und Crook et al., 2008). Peteraf & Barney (2003) argumentieren, dass der Resource-based view durch seine andere Perspektive schlicht andere Phänomene erklärt als der Market-based view und entsprechend beide Theorien eine parallele Existenzberechtigung aufweisen (vergleiche zu diesem Diskurs unter anderem auch Ketchen, Hult & Slater, 2007; Connor, 2007 und Zhou, Li, Zhou & Su, 2008). Der derzeitige Stand der Diskussion spricht dafür, dass auch weiterhin beide Erklärungsmodelle für Überperformance, auch wenn sie regelmäßig falsifiziert wurden, aufgrund ihres in ihrem jeweiligen Feld relativ hohen Erklärungsbeitrages weiter beibehalten und zur Erklärung von Phänomenen der Strategischen Managementforschung herangezogen werden dürften. Gleichzeitig geht das Ringen um eine bessere Theorie zur Erklärung von Performanceunterschieden im Diskurs des Fachgebietes weiter.²⁰

2.1.3 Hypothesen und Hypothesentests

Auch wenn die Forderung nach kritisch-rationalistischen Vorgehensweisen im Theorietest im Rahmen der Strategischen Managementforschung kaum erfüllbar scheint, bleibt doch Poppers Lösung des Hume'schen Induktionsproblems, dass Wissen durch den Ersatz schlechter (falsifizierter) durch bessere Theorien entsteht, als Handlungsmaxime zur Wissensgenerierung im Fachgebiet der Strategischen Managementforschung wegweisend. Zentral für ein tatsächliches (mehr oder weniger gutes) Verständnis des Forschungsgegenstandes ist es damit, Theorien zu entwickeln und diese empirisch zu überprüfen. Theorieentwicklung und Hypothesentest müssen dabei nicht als Einheit verstanden werden, Theorien können durchaus zunächst durch reines „theoretisieren“ entstehen und entwickelt werden, früher oder später müssen sie jedoch unausweichlich dem Test der Realität ausgesetzt werden. Dies geschieht durch die deduktive Prüfung der Theorie. Hierzu werden aus der Theorie Schlussfolgerungen gezogen und Sätze gebildet, diese auf ihre logischen Beziehungen untereinander und zu anderen Sätzen und Theorien untersucht. Schließlich müssen aus den Sätzen einer Theorie Hypothesen zu erwarteten Kausalbeziehungen formuliert und empirisch getestet werden (Popper, 1989, 7). Um sinnvolle Aussagen aus Hypothesen abzuleiten, müssen diese empirisch gehaltvoll sein. Eine Hypothese ist umso gehaltvoller, je mehr Möglichkeiten der Falsifikation sie bietet (Maurer, 2004). Kant (Kant & Eisler, 1930/2009) definiert Hypothesen als Sätze, die „die Verständlichkeit des angenommenen Erklärungsgrundes oder dessen Einheit (ohne Hilfhypothese), die Wahrheit (Übereinstimmung unter sich selbst und mit der Erfahrung der daraus abzuleitenden Folgen), und endlich die Vollständigkeit des Erklärungsgrundes zu ihnen, die auf nichts mehr noch weniger zurückweisen, als in der Hypothese angenommen worden, und das, was a priori synthetisch gedacht war, a posteriori analytisch wieder liefern und dazu zusammenstimmen.“

²⁰ Anzumerken ist allerdings, dass ein Wettstreit zwischen Theorien im Popperschen Sinne immer nur dann möglich ist, wenn diese wiederholt empirischen Überprüfungen ausgesetzt werden. In vielen Gebieten der Strategischen Managementforschung ist dies der Fall, doch beklagen etwa Singh, Ang & Leong (2003) hier an verschiedenen Stellen auch Defizite und rufen zu mehr Forschungen auf, die die empirische Evidenz von Theorien durch Wiederholungsstudien verbessern.

Er teilt, auf Leibniz' Betrachtung zu aus sich selbst heraus evidenten und nichtevidenten Sätzen (Leibniz, 1704/1961, 361) aufbauend, Urteile, die zu Hypothesen gebildet werden, in analytische und synthetische Urteile ein. Bei ersteren geht das Prädikat aus dem Subjekt hervor, die Hypothese ist also in sich immer wahr. Bei letzteren sind Subjekt und Prädikat verschieden. Erkenntnisgewinn ist im Kontext der Sozialwissenschaft nur bei Prüfungen synthetischer Hypothesen zu erwarten, da analytische, ex definitionem, immer wahr sind und damit keine Theorien falsifizieren können (Maurer, 2004)²¹. Ein Beispiel für eine analytische Hypothese ist etwa die Aussage: „der Junggeselle ist unverheiratet“. Hier liefert der zweite Teil der Aussage keine nähere Bestimmung des ersten Teiles, das Prädikat, der Zustand des ledig seins, geht aus dem Subjekt „Junggeselle“ hervor, die Aussage kann logisch nicht falsch sein. Das Gegenteil hiervon sind synthetische Sätze, die wahr oder falsch sein können, etwa: „Mein Bruder ist unverheiratet“ (Beispiele entnommen aus Powell, 2001, 881).

Diese Unterscheidung ist im Kontext der Strategischen Managementforschung insbesondere insofern von Bedeutung, als dass diese häufig mit dem Vorwurf konfrontiert wird, analytische Hypothesen zu prüfen (vergleiche Abschnitt 2.3). Allgemein ist zu fordern, „riskante [= synthetische] Hypothesen statt inhaltsleerer Aussagen“ (Maurer, 2004, 25) zu überprüfen. Diese Forderung ist nicht trivial, da Wissenschaftler durch die Einbindung in ihre sozialen Räume, Systeme und Denkmuster dazu angeregt sein können, die Falsifikation ihrer Hypothesen nicht zu riskieren und stattdessen Hypothesen zu bilden, die nahe am analytischen Typus liegen (ebd.). Damit lassen sich jedoch keine Theorien falsifizieren, Erkenntnisfortschritt kann nicht stattfinden. Dabei ist es keinesfalls notwendig, nur solche Hypothesen zu prüfen, die man zu falsifizieren erwartet. Auch wenn die Verifikation einer Theorie nicht gelingen kann, steigt ihre Erklärungsmacht doch mit zunehmendem „Bewährungsgrad“ (Popper, 1973, 30-33). Theorien werden, auch wenn sie niemals verifiziert werden können, doch durch häufige Nichtfalsifikation der aus ihnen abgeleiteten Hypothesen gestärkt, sofern diese synthetisch sind, also so formuliert sind und unter Bedingungen geprüft werden, dass eine Falsifikation leicht denkbar wäre.

In der Strategischen Managementforschung lässt sich das Problem der analytischen Sätze an aktuellen Diskussionen dazu nachvollziehen, ob es sich bei der Theorie des Resource-based view um ein analytisches oder synthetisches Aussagesystem handelt, dem Tautologievorwurf.²² Im Rahmen des Resource-based view tritt das tautologische Problem beim Konzept des Wettbewerbsvorteils auf, des competitive advantage, der ja, so besagt die Theorie, auf nachhaltig wertschöpfungsrelevante Ressourcen zurückzuführen ist (Barney, 1991; Peteraf, 1993). Aus der Theorie des Resource-based view lässt sich nun für den Zweck des Theorietests die Hypothese ableiten, dass das Vorhandensein von Ressourcen im Sinne des Resource-based view zu einem Wettbewerbsvorteil führt, der wiederum überlegene Performance erklärt. Der Wettbewerbsvorteil lässt sich also aus der Existenz rarer und wertschöpfungsrelevanter Ressourcen heraus beschreiben. Im Zuge des Hypothesentests stellt

²¹ Maurer (S. 16) schreibt die Unterscheidung in analytische und synthetische Sätze Leibniz zu, diese Zuschreibung kann ich nach den mir vorliegenden Primärquellen nicht nachvollziehen. Vielmehr scheint mir diese Unterscheidung, wie geschildert, auf Immanuel Kants Kritik der reinen Vernunft (zitiert nach Kant & Eisler, 1930/2009) zurückzugehen.

²² Eine Tautologie ist ein logischer Zirkelschluss, ein Satz, der unabhängig vom Wahrheitsgehalt seiner Bestandteile für sich immer wahr ist. Tautologische Sätze in Theorien sind also immer analytische Sätze (Popper, 1989).

der Wissenschaftler nun eine wie auch immer definierte Überperformance eines Unternehmens fest und folgert hieraus, dass das oder die besonders erfolgreichen Unternehmen offensichtlich einen Wettbewerbsvorteil gegenüber den Konkurrenten in ihrem Markt aufweisen. Als nächster Schritt wird nun untersucht, inwieweit dieser Wettbewerbsvorteil auf bestimmte Ressourcen zurückgeführt werden kann, beziehungsweise, welche Ressourcen diesen Wettbewerbsvorteil generieren. Findet der Forscher keine Ressource, die er kausal mit dem Wettbewerbsvorteil in Zusammenhang bringen kann, war sein methodisches Konzept offenbar nicht in der Lage, die tatsächlich relevante Ressource aufzuspüren oder deren Relevanz zu belegen. In diesem Gedankengebäude ist es nicht vorstellbar, dass *keine* Ressource existiert, die den Wettbewerbsvorteil erklärt. Das Konzept des Resource-based view ist damit in sich analytisch beziehungsweise tautologisch, es kann nicht falsifiziert werden. Da jeder beobachtete Wettbewerbsvorteil per Definition durch Ressourcen zu erklären ist und Überperformance durch Wettbewerbsvorteile erklärt wird, kann der Resource-based view Überperformance nicht *nicht* erklären. Zwar tut diese Powell (2001) entnommene Fundamentalkritik dem Resource-based view in dieser Generalität Unrecht,²³ doch macht sie das zentrale Problem der Tautologien im Bereich der Strategischen Managementforschung deutlich.

Im Rahmen des in diesem Kapitel geschilderten Wissensschöpfungsprozesses ist, da die Prinzipien der deduktiver Forschungslogik und des instrumentalistischen Theorieverständnisses im Sinne Friedmans als etablierter State of the Art gelten können, die zentrale methodische Fehlerquelle, falls eine solche existieren sollte, im Rahmen des Hypothesentests zu erwarten.

2.1.4 Operationalisierung und Indikatoren

Wesentliche Voraussetzung eines Hypothesentests ist die Operationalisierung der Konstrukte, anhand derer Sätze überprüft werden sollen. Erst dieser Schritt ermöglicht den Test einer Theorie an der Empirie. Durch die Operationalisierung werden Konstrukte in erfassbare Variablen überführt, die gemessen und zueinander in Beziehung gesetzt werden können. Opp (1970) bezeichnet diesen Prozess synonym als „operationales definieren“ und „Spezifikation von Indikatoren für einen Begriff“ (S. 130). Der Prozess des Operationalisierens bedeutet, einem abstrakten Konstrukt beobachtbare Ereignisse, sogenannte Designata, zuzuweisen und gegebenenfalls mehrere Designata untereinander logisch zu verknüpfen. Je nachdem, wie gut die zugewiesenen Designata und ihre Verknüpfungen dem abstrakten Begriff entsprechen, wird die Operationalisierung als valide beziehungsweise weniger valide bezeichnet. Ein Designatum oder ein Bündel an Designata, die in einer operationalen Definition mit einem

²³ Peteraf & Barney (2003) argumentieren zur Verteidigung ihrer Theorie insbesondere, dass die Kausalität zwischen der Existenz von raren Ressourcen und der Generierung eines Wettbewerbsvorteils keineswegs trivial ist, da sich die Ressourcen sowohl auf die Kosten der Gütererstellung als auch auf den Kundennutzen der Güter auswirken können. Es ist demnach denkbar, dass Ressourcen im Sinne des Resource-based view zwar die Qualität der Güter nachhaltig verbessern, aber so kostenintensiv sind, dass sie keinen Wettbewerbsvorteil schaffen. Sirmon, Gove & Hitt (2008) diskutieren die Auswirkungen von Managemententscheidungen zur Umsetzung von Ressourcen in einen Wettbewerbsvorteil im Hinblick auf Reaktionsmuster der Konkurrenz, ähnlich auch im Managerial Rents Model von Castanias & Helfat (2001). Powell & Arregle (2007) ergänzen den ressourcenbasierten Ansatz um einen Blick auf die Fehler, die Unternehmen bzw. Manager machen oder vermeiden und die ebenfalls ein (firmeninternes) Erklärungsmuster für Performance-Differenzen darstellen. Festzuhalten bleibt jedoch, dass der Resource-based view einiger Erweiterungen und konzeptioneller Kunstgriffe bedurfte, um dem Tautologievorwurf zu entgehen.

abstrakten Konstrukt verbunden werden, wird als Indikator dieses Konstrukts bezeichnet (Definition aus Opp, 1970, 131). Der Return on Equity (ROE) beispielsweise beschreibt als Indikator das abstrakte Konstrukt „Leistung“, in dem die Designata „finanzieller Überschuss“ und „Eigenkapital“ in eine definierte Beziehung gesetzt werden, der Periodengewinn wird durch das in der Periode durchschnittlich vorhandene Eigenkapital dividiert. Der ROE ist damit ein durch zwei Designata gebildeter Indikator für das Konstrukt Leistung.

Ein Indikator ist umso valider, je besser er in der Lage ist, das Konstrukt, das er beschreiben soll, abzubilden. Validität bedeutet dabei nicht, dass ein ermitteltes Ergebnis, etwa die Ausprägung eines Indikators, notwendigerweise mit einer objektiven Wahrheit bezüglich des abstrakten Konstrukts übereinstimmt. Die Existenz einer solchen objektiven Wahrheit wird im Bereich der Performance-Messung beispielsweise von Nørreklit, Nørreklit & Mitchell (2007) rundheraus abgelehnt, unbenommen des Problems der Unmöglichkeit der Verifikation wissenschaftlicher Theorien (Popper, 1989). Validität entsteht nach Nørreklit und Kollegen dadurch, dass ein intersubjektives Verständnis von Sinn und Aussage des Methodenspektrums geschaffen wird (vergleiche auch Lamnek, 1995). Die einfachste Form dieser intersubjektiven Akzeptanz stellt die Inhaltsvalidität dar (Combs et al., 2005; vergleiche auch Popper, 1989, 18-21). Diese wird erreicht, wenn Übereinstimmung herrscht, dass ein bestimmtes Konstrukt in den Bereich fällt, in den es fallen soll, also dass etwa eine bestimmte Leistungsoperationalisierung tatsächlich Leistung misst. Validität wird im Weiteren dadurch gewonnen, dass eine Messgröße einen Bezug zur empirischen Realität herstellen, also die (angenommenermaßen existente) reale Welt abbilden kann (Nørreklit et al., 2007). Dieser Grundforderung nähert man sich durch Operationalisierungen an, die den Anforderungen konvergenter und diskriminanter Validität genügen (Combs et al., 2005). Konvergente Validität ist erreicht, wenn zwei Operationalisierungen desselben Konstrukts hoch korreliert sind, diskriminante Validität liegt vor, wenn Operationalisierungen, die unterschiedliche Konstrukte messen sollen, wenig oder nicht miteinander korrelieren.²⁴

Im englischen Sprachraum, in dem die relevante Literatur im Fachgebiet der Strategischen Managementforschung publiziert wird (vergleiche Abschnitt 4.1.3), wird in aller Regel auf eine begriffliche Trennung zwischen dem Prozess des Operationalisierens, also des Zuweisens von Indikatoren zu einem Konstrukt, und dem Indikator selbst verzichtet (vergleiche etwa Crook, Ketchen, Combs & Todd, 2008; Boyd, Gove & Hitt, 2005; Holloway, Lewis & Mallory, 1995; Venkatraman & Ramanujam, 1986). Stattdessen wird generell von „measurement“ als dem Prozess der Erfassung des Konstrukts und von „measure“ im Sinne eines Indikators gesprochen. Dieser Vorgehensweise schließe ich mich in der vorliegenden Arbeit an und verwende den Begriff der „Operationalisierung“ als Synonym zum Begriff des „Indikators“. ROE, im obigen Beispiel, ist in der Sprachregelung dieser Arbeit also eine Operationalisierung des abstrakten Konstrukts Leistung.

²⁴ Sozialkonstruktivistisch betrachtet kann von der Annahme einer existenten realen Welt auch abgegangen werden, Validität entsteht dann schlicht durch Akzeptanz einer gemeinsam konstruierten Wirklichkeit in einer Gruppe (vergleiche hierzu Denzau & North, 1994), diskriminante und konvergente Validität beziehen sich dann auch entsprechend auf in der Gruppe definierte Konstrukte. Um die Klarheit der Argumentation zu gewährleisten verwende ich im Folgenden kritisch-rationalistische Argumentationen mit der Annahme der Existenz einer äußeren Realität (vergleiche auch: Flach, 2008; Lehmann-Waffenschmidt, 2006). Grundsätzlich ist die Definition und Operationalisierung der Leistung von Organisationen allerdings ein hervorragendes Beispiel für den sozialkonstruktivistischen Charakter von Konstrukten in den Sozialwissenschaften.

Die Leistung einer Organisation ist das zentrale Konstrukt in der Strategischen Managementforschung, Indikatoren, die dieses Konstrukt erfassen sollen, stellen die wichtigsten abhängigen Variablen im Fachgebiet dar.²⁵ Die Operationalisierung des Konstrukts Leistung geht jedoch mit verschiedenen Schwierigkeiten einher. Bereits die Ambiguität des Leistungsbegriffs und damit des Leistungskonstrukts ist eine erste wesentliche Hürde. Die Definition des Leistungsbegriffs ist hochgradig komplex, wie im Folgenden zu diskutieren sein wird.

2.2 Der Leistungsbegriff

2.2.1 Leistung – Das Definitionsproblem

Voranzustellen ist, dass sich die vorliegende Arbeit ausschließlich mit der Leistung von Organisationen, nicht mit der von Individuen oder (Arbeits-)Gruppen befasst. Zur Betrachtung der Möglichkeiten und Grenzen der individuellen Leistungsmessung am Arbeitsplatz oder der Erfassung von Leistungen von in Organisationen eingebetteten kleineren Gruppen und Teams sei beispielsweise auf Holloway, Lewis & Mallory (1995), Thommen & Achleitner (2006), Schomann (2001) oder Marginson & McAulay (2008) verwiesen.

Die erste Schwierigkeit auf dem Weg zur Messung der Leistung einer Organisation besteht bereits im Begriff selbst. Beginnen wir die Betrachtung mit einem Blick in das Lexikon. Meyers Universallexikon definiert Leistung als den „Grad der körperlichen und geistigen Beanspruchung sowie deren Ergebnis“. ²⁶ Leistung besteht demnach aus zwei Komponenten: Dem Input, also dem Einsatz, der in eine bestimmte Tätigkeit oder einen bestimmten Prozess investiert wird, und dem Output, also dem Ergebnis dieser Tätigkeit oder dieses Prozesses. Im Rahmen der Betrachtung der Leistung von Organisationen definiert Meyers die betriebswirtschaftliche Leistung als: „Die Menge oder der Wert, der innerhalb eines Zeitraumes im betrieblichen Produktionsprozess hervorgebrachten Sachgüter oder der bereitgestellten Dienstleistungen.“

Diese Definition taugt zur Beschreibung der Leistung von Organisationen allgemein insofern, als dass im weiteren Sinne alle erreichten Ziele einer Organisation, sei es der Jahresüberschuss eines Unternehmens, die Brandbekämpfung einer Feuerwehr, die Versorgung von Hilfsbedürftigen in karitativen Einrichtungen oder das Ergebnis einer Mannschaft im Leistungssport im weitesten Sinne als Sachgut oder Dienstleistung, zumindest jedoch als Leistungsergebnis fassen lassen. Diese Definition des Leistungsbegriffs deckt sich jedoch kaum mit dem semantischen Inhalt des Begriffes „Leistung“ im normalen deutschen Sprachgebrauch.

Becker bezeichnet in seinem Standardwerk zur betrieblichen Leistungsmessung die Verwendung des Begriffes „Leistung“ im allgemeinen Sprachgebrauch als „Leerformel“ (Becker, 2003, 2f) und nennt den Terminus folgerichtig inhaltsleer. In der Tat spiegelt sich in Begriffen wie „Leistungsprinzip“²⁷, „gutes Geld für gute Leistung“²⁸ und anderen mehr, denen wir uns im

²⁵ Vergleiche hierzu die Literaturreview in Kapitel 4.

²⁶ vgl. lexikon.meyers.de

²⁷ Eine häufig genutzte Vokabel etwa in Reden der derzeitigen Bundeskanzlerin Dr. Angela Merkel, vgl. <http://www.bundestkanzlerin.de>.

²⁸ Etwa im derzeitigen Diskurs um die Höhe von Managergehältern, vgl. bspw. die Debatte unter diesem Schlagwort im Rahmen der Online-Präsentation der Wochenzeitung „Die Zeit“, www.zeit.de.

täglichen Leben gegenübersehen, das zentrale Problem des Leistungsbegriffs wieder: Eine erhebliche begriffliche Unschärfe. Was etwa ist der Maßstab für „gute Leistung“ eines Managers, die dann gutes Geld wert wäre? Die Aktienkursentwicklung des von ihm geleiteten Unternehmens? Die Zahl der geschaffenen Arbeitsplätze? Das Überleben des Unternehmens? Eine gute Umweltbilanz? Corporate Social Responsibility? Während für den öffentlichen Diskurs ein schwammiges Verständnis des Leistungsbegriffs unter Umständen noch ausreichen mag ist klar, dass, wenn der Begriff „Leistung“ in einem operativ nutzbaren Sinne im Rahmen wissenschaftlicher Arbeit, etwa in Hypothesentests, gebraucht werden soll, eine saubere Definition unverzichtbar ist.

Dabei kennt bereits die Betriebswirtschaftslehre verschiedene Verwendungen des Leistungsbegriffs. Zur Illustration der Problematik vergleiche ich im Folgenden Definitionen in drei Standard-Lehrbüchern. Im Kontext der Kosten- und Leistungsrechnung, also dem Rechnungswesen, nennt man Leistung „den – in Geld ausgedrückten – sachzielbezogenen Wertezugang im Rahmen der ordentlichen Geschäftstätigkeit in einer Periode“ (Schmalen, 2002, 642). Thommen & Achleitner (2006) verstehen unter der Leistung eines Unternehmens seinen Erfolg, messbar anhand verschiedener Größen wie Rentabilitäten oder Wirtschaftlichkeit. Bea, Friedl & Schweitzer (2006) betrachten den Leistungsprozess eines Unternehmens und definieren diesen als „die Beschaffung von Dienstleistungen und Sachgütern (Produktionsfaktoren), die Kombination (Transformation) von Produktionsfaktoren zu Wiedereinsatz- und Absatzgütern sowie den Absatz von Gütern“ (S. 1). Die Leistung eines Unternehmens ist demnach das Ergebnis dieses Leistungsprozesses. Der Unterschied ist klar: Während in der am Rechnungswesen orientierten Definition den Wertzugang misst, findet sich in der erfolgsorientierten Definition ein Bezug zwischen Aufwand und Kosten. Die prozessorientierte Ansicht schließlich betrachtet überdies auch die Tätigkeit, den Weg, der zu einem bestimmten Ergebnis führt. Das Problem der semantischen Mehrfachbesetzung des Leistungsbegriffs ist auch keineswegs auf den deutschen Sprachraum begrenzt. Im englischen Sprachgebrauch umgibt den entsprechenden Begriff „Performance“ eine ähnliche Bedeutungsambiguität, sowohl hinsichtlich der Vielfalt der unter Performance gefassten Begrifflichkeiten als auch im Hinblick auf die unklare Spezifikation hinsichtlich Prozess- oder Ergebnisdimension (diskutiert etwa in Lebas & Euske, 2007).

In diesem Kontext ist die Diskussion um eine Differenzierung der Termini Leistung und Wettbewerbsvorteil (im englischen Sprachraum *performance* und *competitive advantage*) erhellend, wie sie insbesondere im Diskursraum des Resource-based view (vgl. Barney, 1991; Peteraf, 1993) stattfindet und bereits in Abschnitt 2.1 vor dem Hintergrund der Problematik analytischer Sätze angerissen wurde. Der Theorie des Resource-based view folgend führen Ressourcen zu einem Wettbewerbsvorteil, der wiederum zu Leistung führt. Diese Kausalität konnte empirisch mehrfach bestätigt werden (vgl. etwa Crook et al., 2008 oder Newbert, 2008 als Beispiele für neuere Untersuchungen hierzu). Fraglich ist, ob es sich bei Leistung und Wettbewerbsvorteil tatsächlich um zwei getrennte Konzepte handelt oder ob diese beiden Konzepte so nahe beieinanderliegen, dass eine Unterscheidung unsinnig, bzw. tautologisch wird. Von Wiggins & Ruefli (2002) etwa werden beide Begriffe synonym verwendet. Auch andere Autoren kommen zu dem Schluss, dass die Trennung der Begriffe zumindest unscharf ist: „Within many studies, the term *competitive advantage* is almost synonymous with *performance* in the sense that *competitive advantage* is generally used to describe the relative

performance of rivals in a given (product) market environment" (Peteraf & Barney, 2003, 313; ausführlich auch in Grahovac & Miller, 2009 und Durand & Vaara, 2009).

Eine tiefere Auseinandersetzung mit dieser Thematik findet sich bei Powell (2001). Der logische Schluss, dass Wettbewerbsvorteile immer zu einer in Relation zur Leistung der Konkurrenz überlegenen eigenen Leistung führen, ist im Umkehrschluss nicht ebenso wahr: Es kann nicht ausgeschlossen werden, dass es Unternehmen mit überlegener Leistung gibt, die keine Wettbewerbsvorteile aufweisen, daher können die beiden Begriffe nicht synonym verwendet werden. Überdies gibt es auch Unternehmen, die ihren eigentlich vorhandenen Wettbewerbsvorteil nicht in Performance umsetzen können. Entsprechend handelt es sich bei Wettbewerbsvorteil und Leistung um zwei getrennte Konzepte (vergleiche auch Newbert (2008); Peteraf & Barney (2003); Powell (2001)). „Wettbewerbsvorteil“ bezieht sich auf Strategien und Fähigkeiten des Unternehmens, die den Strategien und Fähigkeiten der Konkurrenz überlegen sind, während die „Leistung“ in diesem Kontext als die hierdurch generierte Rente definiert wird (Newbert, 2008). Im Interpretationsschema Beckers (Becker, 2003) entspräche der „competitive advantage“ damit also der Tätigkeitskomponente der Leistung, während die „performance“ als Output-Komponente das Ergebnis des Leistungserstellungsprozesses darstellt. Diese Unterteilung erscheint methodisch zielführend, bricht sie doch nun die sprachliche Mehrfachbelegung des Leistungsbegriffs auf: In eine tätigkeits- und eine ergebnisorientierte Komponente. Diese Trennung ist, wie bereits oben erwähnt, zentral für operatives Arbeiten in Hypothesentest im Feld der Strategischen Managementforschung. Für die Zwecke der weiteren Untersuchung definiere ich Leistung, obigen Überlegungen folgend, daher folgendermaßen:

Die Leistung einer Organisation ist das Ergebnis der Summe ihrer Tätigkeiten.

Diese Definition ist sehr allgemein und rein ergebnisorientiert. Leistung im Sinne der Strategischen Managementforschung ist damit kein Prozess und trifft keine Aussagen über den Input im Sinne der Meyer'schen Lexikondefinition, sondern stellt nur und ausschließlich das Endergebnis des Leistungserstellungsprozesses dar.

Aufgrund des Metastudiencharakters kann und darf diese Arbeit naturgemäß keine engere Definition von Leistung vorgeben, zumal die wirtschaftsethisch problematische Fragestellung, welche Kenngrößen denn „richtige“ oder ethisch erstrebenswerte Ziele für Unternehmen darstellen (sollen), nicht Gegenstand dieser Arbeit ist. Vielmehr soll sich die Definition auf ein Grundverständnis beschränken. Die Definition, was jeweils spezifisch unter der Leistung der untersuchten Organisationen zu verstehen ist, ist integral mit der jeweiligen Methodik der untersuchten Studien verwoben und wird, explizit oder implizit, vom jeweiligen Untersuchenden definiert. Eine gute Leistungsoperationalisierung in der Strategischen Managementforschung ist damit eine, die im Sinne des in Abschnitt 2.1 diskutierten Verständnisses von Wissenszuwachs in den Sozial- und Geisteswissenschaften zum Test von Hypothesen und damit Theorien funktional und geeignet ist, das Konstrukt Leistung und Kausalitäten in Bezug zu diesem Konstrukt zu erfassen. Insbesondere letzteres ist unter Umständen problematisch, wie die Ausführungen in Abschnitt 4.3 zeigen werden.

Der Leistungsbegriff lässt sich nichtsdestotrotz anhand der folgenden Beobachtungen zur Verwendung des Terminus' weiter eingrenzen:

- *Leistung als subjektives Konstrukt*: Leistung ist immer ein Konstrukt, das funktional zur Bewertung einer bestimmten Situation oder einer begrenzten Menge von Situationen ist. Entsprechend ist es auch nicht möglich, Leistung allgemein zu definieren. Die spezifische Definition des jeweils verwendeten Leistungsbegriffs hängt überdies immer von der subjektiven Bewertung des oder der beteiligten Personen (Untersucher und Untersucher) ab. Umso mehr ist eine klare Kommunikation wichtig, was in einer spezifischen Situation von einer bestimmten Personen bzw. einem Personenkreis unter „Leistung“ genau verstanden wird.²⁹
- *Leistung als relatives Konstrukt*: Leistung existiert nicht im maßstabslosem Raum sondern bedarf, wenn ein entsprechender Messwert eine Aussage tragen soll, immer einer wertenden Komponente (Lebas & Euske, 2007; Pike & Roos, 2007). Diese kann implizit oder auch expliziert vorliegen und drückt sich im täglichen Sprachgebrauch beispielsweise durch Begriffserweiterungen wie „Normalleistung“, „Hochleistung“, „gute Leistung“, „schlechte Leistung“ und so weiter aus (Pawlowsky & Mistele, 2008; Pawlowsky, 2008). Notwendig hierfür ist ein Bewertungsmaßstab, der wiederum explizit oder implizit vorliegen kann. Während im normalen Sprachgebrauch ein implizites Maß („gute Leistung“) ausreichen mag, ist es im ökonomischen und mehr noch im wissenschaftlichen Bereich unumgänglich, den Bewertungsmaßstab offen zu legen, an dem Leistung gemessen werden soll. Dieser kann ein Absolutwert sein, beispielsweise die Leistung als das Überleben eines Unternehmens über einen bestimmten Zeitraum oder das Erreichen eines Gewinnziels in einer spezifischen Periode, oder auch ein relativer Wert, etwa die Geschäftsentwicklung im Vergleich zu internen oder externen Vorgaben oder Benchmarks oder die Entwicklung des Aktienkurses eines Unternehmens in Relation zu einem relevanten Marktindex oder einem trend- und risikoangepassten Normalwert.³⁰
- *Leistung als mess- und bewertbares Konstrukt*: Der subjektive und relative Charakter von Leistung beinhaltet implizit auch, dass der Leistungsbegriff nur dann mit Inhalt gefüllt werden kann, wenn die Leistung in irgendeiner Form valider Messung oder Bewertung zugänglich ist. Hierbei kann es sich, wie zumeist der Fall, um eine quantifizierte Art der Messung handeln, etwa die Ermittlung von Rentabilitätszahlen oder Produktivität, aber durchaus auch um qualitative Einschätzungen, etwa im Rahmen der Bewertung der Leistungsfähigkeit einer Organisation durch Expertenurteile.

Diese Beschreibung des Leistungsbegriffs trifft bewusst keine Aussage über den Typus an Organisation, der die Leistung erbringt. Das Interesse der Strategischen Managementforschung richtet sich primär auf Unternehmen, doch selbstverständlich kann und muss auch die Leistung von nichterwerblich orientierten Organisationen erfasst und gemessen werden und steht damit Fragestellungen, die in dieser Arbeit diskutiert werden, als abhängige

²⁹ Vergleiche zur Thematik externer und interner Validität auch Lamnek (1995).

³⁰ Lebas & Euske (2007) gehen noch einen Schritt weiter und betonen, dass Leistung nur aus der subjektiven Wahrnehmungskonstruktion der Bewertenden verstanden werden kann. Nørreklit, Nørreklit & Mitchell (2007) betonen ebenfalls die sozialkonstruktivistische Dimension des Leistungsbegriffs. Dieser Argumentation soll an dieser Stelle jedoch nicht gefolgt werden, um Inkommensurabilitätsprobleme zu vermeiden. Verständnisleitend bleibt das in Abschnitt 2.1 entwickelte kritisch-rationalistische Wissenschaftsverständnis. Vergleiche zur Diskussion von Validität in der Leistungsmessung Steigenberger, Pawlowsky, Lübcke & Schulte (2010).

Variable zur Verfügung. Aus Gründen der Fokussierung und der Vermeidung allzu häufiger Verzweigungen im Argumentationsgang bezieht sich die folgende Untersuchung jedoch ausschließlich auf erwerbswirtschaftlich orientierte Organisationen, also Unternehmen. Die Begriffe Organisation und Unternehmen werden daher im Weiteren synonym verwendet.

Zumeist bezieht sich das Konstrukt der Leistung auf gegenwärtige oder vergangene Ereignisse. Allerdings kann man die Leistung einer Organisation durchaus auch aus einer potentialorientierten Sichtweise betrachten (Meyer, 2007; Lebas, 1995). Meyer (2007) etwa definiert Leistung als die abgezinsten zukünftigen cash-flows eines Unternehmens, also als eine Erwartung über zukünftig zu generierende Renten. Dieser Ansatz wirft jedoch ein systematisches Operationalisierungsproblem auf, sind doch zukünftige Erträge ad definitionem vom heutigen Zeitpunkt aus nicht messbar und Kausalaussagen daher kaum zugänglich. Häufig wird ein vergangenheitsbezogener Leistungsbegriff vollauf genügen, in vielen Fällen kann es aber durchaus notwendig werden, sich dem Leistungsverständnis als Potential zur Generierung von Renten in der näheren oder mittleren Zukunft zumindest anzunähern. Damit kann die retrospektive Orientierung des Leistungsbegriffs durchaus als eine Schwäche der Operationalisierung des Konstrukts Leistung verstanden werden. Auf das Problem der zeitlichen Orientierung der Konstrukte wird im weiteren Gang dieser Arbeit in Abschnitt 4.4 noch detaillierter einzugehen sein.

2.2.2 Business Performance Measurement – Abgrenzung

Eine wesentliche Abgrenzung muss zwischen dem Problem der Performanceoperationalisierung, wie sie in der Strategischen Managementforschung notwendig und Gegenstand dieser Arbeit ist, sowie dem Business Performance Measurement erfolgen. Die Messung der Leistung ist im ökonomischen Raum in einer Vielzahl unterschiedlicher Konstellation relevant. So muss die Arbeit von Mitarbeitern bewertet werden, die Leistung von Arbeitsgruppen und Abteilungen will eingeschätzt werden, Unternehmenslenker verlangen nach belastbaren, zumeist quantitativen Messgrößen sowohl finanzieller als auch nichtfinanzieller Art, auf denen sie ihre Entscheidungen basieren können (Eccles, 1995). All dies sind Aufgabenstellungen des Business Performance Measurement, der Bereitstellung von Leistungsindikatoren aus einem Unternehmen zur Steuerung dieses Unternehmens oder zur Unternehmensanalyse. Business Performance Measurement befasst sich allgemein mit der Bereitstellung von Daten zur Information von Entscheidungsträgern, zumeist innerhalb der in Frage stehenden Organisation (etwa Managementinformationssysteme), aber auch aus dem Kreis externer Stakeholder³¹ (etwa zur Information von potentiellen oder aktuellen Investoren). Durch die Messwerte sollen Entscheidungen unterstützt werden, die in einem Zusammenhang mit dem Objekt stehen, auf das sich die Messwerte beziehen (Pike & Roos, 2007). Entscheidend ist hierbei, wie ebenfalls von diesen beiden Autoren angemerkt wird, ein vollständiges Verständnis und eine saubere Definition und Abgrenzung dessen, was gemessen werden soll. Aus einer deskriptiven Perspektive heraus ermöglichen Business Performance Measurement Systeme die Messung und Quantifizierung der Effizienz und Effektivität von Handlungen (vergangenheits-

³¹ Stakeholder eines Unternehmens sind alle Personen oder Gruppen, die ein begründetes Interesse an diesem Unternehmen haben, diesem also durch explizite oder implizite Verträge verbunden sind. Vergleiche ausführlich hierzu Abschnitt 4.3.1 sowie zur Diskussion um Definitionen des Stakeholder-Begriffs Mitchell, Agle & Wood, 1997.

bezogene Perspektive), aus einer strategisch-handlungsorientierten Sichtweise heraus helfen Messsysteme im Rahmen der Implementierung von Strategien, indem Hauptziele für Untergruppen von Organisationen heruntergebrochen und diese in ihrem Zielkanon so auf Linie mit dem strategischen Hauptziel gebracht werden, sowie bei der Überprüfung und Über-/Erarbeitung von Strategien (zukunftsbezogene Perspektive, vergleiche auch Franco-Santos et al., 2007).

Ein wesentliches Kennzeichen von Business Performance Measurement Systemen ist, dass sie von Akteuren erstellt werden, die in der in Frage stehenden Organisation verortet sind und über entsprechende Möglichkeiten der Informationsgewinnung verfügen. Auch wissenschaftliche Fragestellungen können auf den Methoden und Kennziffern des Business Performance Measurement basieren, wenn es den Autoren gelingt, ebenfalls eine derartige Insider-Perspektive einzunehmen. Im Rahmen der Strategischen Managementforschung ist dies bei Untersuchungen der Fall, die mit sehr geringen Fallzahlen arbeiten oder als reine Fallstudien konzipiert sind (einige neuere Beispiele: Ozcan & Eisenhardt, 2009; Capaldo, 2007; Burgelman & Grove, 2007; Haas & Hansen, 2007). In derartigen Forschungsansätzen geht es darum, die jeweilige Organisation oder Personen beziehungsweise Gruppen von Organisationen oder Personen in der Tiefe zu durchleuchten, um einen Zusammenhang zwischen deren Erfolg (oder Misserfolg) und spezifischen manifesten oder latenten Kontextvariablen, seien sie interner (etwa Prozesse, angewandte Strategien, Erwartungen) oder externer Art (etwa Umweltveränderungen, tatsächlicher oder potentieller Wettbewerbsdruck), herauszuarbeiten. Diese Herangehensweise zeichnet sich methodisch insbesondere durch die folgenden Rahmenbedingungen aus:

- Die Zahl der zu untersuchenden Objekte ist relativ klein, der Zugang zu den benötigten Informationen ist so verhältnismäßig einfach möglich, da die Informationssammlung im Einzelfall relativ umfangreich ausfallen kann.
- Bereits zu Beginn der Untersuchung ist bekannt, welche Objekte (Organisationen oder Mitarbeiter) betrachtet werden sollen. Es besteht also kein Auswahlproblem.
- Die abhängige Variable „Leistung“ kann einzelfallbezogen operationalisiert und interpretiert werden.

In der Strategischen Managementforschung liegen zumeist jedoch völlig andere Problemlagen vor. Regelmäßig ist es notwendig, den Untersuchungshorizont von der Fallstudie hin auf ein größeres Sample zu verbreitern, um valide Hypothesentests möglich zu machen. Dies bedeutet, dass die Leistung mit Rücksicht auf die notwendige Sammlung von Daten operationalisiert werden muss. Diese Datensammlung muss mit einem vertretbaren Aufwand bei einer relativ großen Zahl an Organisationen möglich sein. Zudem ist es notwendig, alle Organisationen im Sample mit der gleichen Operationalisierung zu erfassen, einzelfallbezogene Operationalisierungen sind also nicht möglich. Dies hat zur Folge, dass das Instrumentarium des Business Performance Measurement in der Strategischen Managementforschung im Allgemeinen nicht anwendbar ist.

Grundsätzlich sind in den Teilen der Strategischen Managementforschung, die sich mit der Leistung von Organisationen als abhängiger Variablen befassen, im Rahmen von Hypothesentests mit Bezug zur Operationalisierung von Leistung zwei wesentliche Probleme zu lösen: Zum einen die Entwicklung, beziehungsweise Auswahl eines brauchbaren Leistungsindikators,

der sich auf eine Vielzahl von Organisationen mit vertretbarem Aufwand anwenden lässt und gleichzeitig eine befriedigende Konstruktvalidität aufweist, zum zweiten die Eingrenzung der Grundgesamtheit auf eine Größe, die diesem Indikator zugänglich ist, die Populationsselektion. Beide Aufgabenfelder beeinflussen sich gegenseitig und haben einen maßgeblichen Einfluss darauf, ob tatsächlich ein wirkungsvoller Hypothesentest im Popperschen Sinne durchgeführt wird, beziehungsweise mit einem bestimmten Untersuchungsdesign durchgeführt werden kann, wie die weiteren Ausführungen in dieser Arbeit zeigen werden. Die folgenden Betrachtungen zeigen jedoch auch, dass hierbei regelmäßig Schwierigkeiten auftreten.

Im weiteren Fortgang dieser Arbeit richtet sich das Augenmerk auf Forschungen, die mit relativ großen Samples arbeiten. Fallstudienorientierte Arbeiten, für die das Instrumentarium des Business Performance Measurement von Interesse ist, werden aufgrund ihrer geringen praktischen Relevanz in der Strategischen Managementforschung (vergleiche Kapitel 4) im Zuge der folgenden Methodendiskussion nicht mehr explizit untersucht. Viele der im Fortgang der Arbeit diskutierten Probleme sind jedoch auch für Fallstudienanalysen relevant. Die Analyse aktueller Ansätze des Business Performance Measurement zum Zwecke der Bereitstellung von Kenngrößen als Entscheidungsgrundlage für interne und externe Stakeholder ist ebenfalls nicht Gegenstand dieser Arbeit. Es sei hier auf Neely, Kennerley & Adams (2007), Holloway et al., (1995), Schomann (2001), Becker (2003), Meyer (2007), Marginson & McAulay (2008) und insbesondere Franco-Santos et al. (2007) verwiesen.

2.3 Strategische Managementforschung

Die Strategische Managementforschung³² befasst sich mit der Untersuchung von Determinanten, die zu Erfolg oder Misserfolg von Organisationen führen und der Erarbeitung von Empfehlungen für Manager, wie der Erfolg von Unternehmen gegebenenfalls gesteigert werden kann (Combs et al., 2005, 260). Bromiley, 2005, 2, definiert drei Kernaufgaben der Strategischen Managementforschung:

- Die Erklärung des Verhaltens von Unternehmen auf der strategischen Ebene.
- Die Erklärung von Performance-Unterschieden zwischen Unternehmen.
- Die Unterbreitung von Vorschlägen zur Verbesserung der Performance von Unternehmen (normative Aufgabe).

Typische Untersuchungen der strategischen Managementforschung befassen sich etwa mit den Auswirkungen bestimmter Management-Praktiken, Strukturen, Zielsetzungen oder Verhaltensweisen gegenüber Konkurrenten oder Partnern und deren Zusammenhang mit der Leistung der untersuchten Organisation oder der Generierung und dem Einsatz wertschöpfungsrelevanter Ressourcen. Erkenntnisobjekt der Strategischen Managementforschung sind in der Regel erwerbswirtschaftlich orientierte Organisationen, also Unternehmen. Die Strategische Managementforschung zieht ihre Relevanz insbesondere aus der praktischen Bedeutung ihrer Erkenntnisobjekte und ihrer Erkenntnisse für das ökonomische Leben und kann

³² Der folgende Abschnitt basiert wesentlich auf den Ausführungen von Rumelt, Schendel & Teece (1994) und Combs, Crook & Shook (2005). Auf eine Darstellung der historischen Entwicklung des Fachgebietes wird an dieser Stelle aus Fokussierungsgründen verzichtet, der interessierte Leser findet eine solche etwa bei Nerur, Rasheed & Natarajan (2008).

sowohl als Erfahrungs- als auch als normativ orientierte Handlungswissenschaft verstanden werden. Ein Charakteristikum der Strategischen Managementforschung ist überdies auch ihr zumeist retrospektiver Charakter. Der Ausgangspunkt, die Beobachtung eines bereits sichtbaren Ergebnisses, etwa einer Überperformance, erlaubt eine selektive Wahrnehmung der Realität, die zu diesem Punkt geführt hat (Powell, 2001), und damit einen fokussierten Erkenntnisgewinn. Enge thematische Verbindungen bestehen insbesondere zur Organisationssoziologie, aber auch zu Teilen der Mikroökonomie, der Organisationspsychologie und den Politikwissenschaften. Bei der Strategischen Managementforschung handelt es sich um ein Querschnittsfach, dessen Abgrenzung zu verwandten Gebieten nicht immer eindeutig ist.³³

Ein wesentlicher Ansatz der Strategischen Managementforschung ist die Verknüpfung einer oder mehrerer Variablen mit der Leistung einer Organisation, um so den Prozess der Entstehung eines bestimmten Ergebnisses oder den Einfluss bestimmter Strukturen, Bedingungen oder Handlungen auf dieses verstehen zu können. Überdies fallen jedoch auch zahlreiche weitere Fragestellungen in diesen Bereich, die sich mit Kausalitäten auf der Ebene der Gesamtorganisation befassen, beispielsweise Einflussfaktoren auf Strategieentwicklungen oder Strategische Gruppen (etwa Short et al., 2007) oder den Zusammenhang zwischen verschiedenen Managementbedingungen und organisationalem Verhalten (etwa Goranova, Alessandri, Brandes & Dharwadkar (2007), eine Untersuchung zum Zusammenhang zwischen Risikoneigung und Management Ownership). Auch Analysen des Handelns der Unternehmensführung, also des Managementteams, werden zumeist in den Bereich der Strategischen Managementforschung gezählt, die Abgrenzung zu benachbarten Disziplinen ist jedoch allgemein wenig trennscharf (und für die Zwecke dieser Arbeit auch wenig relevant).

Das Konstrukt der Leistung einer Organisation steht zumeist mittelbar oder unmittelbar im Fokus von Theorien und Hypothesentests in der Strategischen Managementforschung und kann, wie in den folgenden Teilen der Arbeit klar werden wird, auf vielerlei Art gemessen werden. Letztgültiges Leistungskriterium einer Organisation bleibt ihr Überleben, beziehungsweise ihre Fähigkeit, zumindest die zum Selbsterhalt notwendige Rendite zu erwirtschaften (Rumelt, Schendel & Teece, 1994). Wesentlich ist nun die Beobachtung, dass es immer wieder Organisationen gibt, die substantiell bessere oder schlechtere Leistungen erbringen als ihre Vergleichsgruppe (vergleiche beispielsweise Rumelt et al. (1991), Teece, Pisano & Shuen (1997) und Wiggins & Ruefli (2002)). Entscheidend ist dabei die Nachhaltigkeit der Über- oder Unterperformance, das heißt die Beständigkeit über eine gewisse Zeit. Das neoklassische Standardmodell kennt ebenfalls Überprofite, allerdings werden diese in kürzester Zeit von konkurrierenden Unternehmen oder in den Überrenten generierenden Markt eintretenden Wettbewerbern aufgezehrt, so dass sich erneut ein Gleichgewicht einstellt, in dem alle Marktteilnehmer annähernd die gleiche Rente generieren, die dem marginalen Unternehmen, also dem Unternehmen, das die ungünstigste Kosten-Umsatz-Struktur aufweist und nicht vom Markt verschwindet, gerade noch das Überleben ermöglicht (Schöler, 2004). Substantielle Unterperformance führt zu einem Ausscheiden aus dem Markt.³⁴ In diesem Modell

³³ Vergleiche ausführlicher zum Abgrenzungsproblem des Fachgebietes auch Nag, Hambrick & Chen (2007). Die Autoren entwickeln aus einer Literaturreview heraus auch weitere Vorschläge zur Definition des Fachgebietes.

³⁴ Diese Vorhersage der Nullprofite ist zwar als allgemeine Aussage aufgrund der beobachtbaren nachhaltigen Überperformance in manchen Bereichen nicht haltbar (Rumelt, Schendel & Teece, 1991), wurde aber von Wiggins & Ruefli (2002) für einen wesentlichen Teil der Unternehmen in vielen Bran-

der vollständigen Konkurrenz bestehen zwischen Unternehmen keine dauerhaften Unterschiede, sie sind in ihren Fähigkeiten, ihrer Ressourcenausstattung usw. identisch. Diese axiomatische Annahme entbehrt naturgemäß logischer und empirischer Fundierung (Powell, 2001), erklärt aber zufriedenstellend wesentliche Teile des Marktgeschehens (vergleiche auch Abschnitt 2.1.2 und Friedman, 1966/1994). Nicht jedoch erklärt diese Theorie das Problem der persistenten Überprofite (Rumelt et al., 1991 und Rumelt et al., 1994, 540-543), also über einen längeren Zeitraum bestehender Performanceunterschiede zwischen Organisationen. Hier setzt die Strategische Managementforschung an.

Methodisch wird häufig von der Beobachtung der abhängigen Variablen Leistung ausgegangen. Auf Grundlage von zur Erklärung geeignet scheinender Theorien oder Ansichten werden Hypothesen zu den Ursachen der Überperformance beziehungsweise zum Zusammenhang dieser Hypothesen mit der unabhängigen Leistungsvariablen generiert und empirisch oder im Gedankenexperiment geprüft. Somit entstehen Erklärungsmodelle, die Leistungsunterschiede zwischen Organisationen erklären können sollen beziehungsweise die Einflüsse bestimmter Strukturen, Handlungen oder Bedingungen auf die Leistung einer Organisation verallgemeinernd darstellen. Bis dato war die Strategische Managementforschung jedoch nicht in der Lage, eine dominante Theorie zu entwickeln, die nachhaltige Performance-Unterschiede erklären könnte (Wiggins & Ruefli, 2002).

Die Suche nach kausalen Zusammenhängen zwischen der Performance einer Organisation und Faktoren, die diese Performance erklären, sieht sich durchaus scharfer Kritik ausgesetzt, die ich in Kapitel 1.1 bereits angerissen habe. Die zentralen Kritikpunkte betreffen die Annahme, es könnte verallgemeinerbare Kausalbeziehungen zwischen bestimmten Determinanten und der Leistung von Organisationen in einer sich stetig verändernden, hoch komplexen Welt grundsätzlich geben (March & Sutton, 1997; auch Gadamer, 1990, vgl. Anmerkungen zur Organisationskybernetik unten). Entsprechend steht der Vorwurf im Raum, dass sich das Fachgebiet aufgrund dieser Unfähigkeit auf tautologische Sätze zurückziehen muss und so nur den Anschein zu erwecken versucht, synthetische Theorien generieren zu können (Nicolai & Kieser, 2002)³⁵, oder durch die Verwendung mathematischer Modelle, die den komplexen,

chen näherungsweise bestätigt und sollte daher in ihrem Erklärungsgehalt nicht gänzlich abgetan werden.

³⁵ Die Kritik von Nicolai und Kieser richtet sich nicht *expressis verbis* gegen die Strategische Managementforschung sondern vielmehr gegen die sogenannte „Erfolgsfaktorenforschung“. Diese wiederum ist dadurch gekennzeichnet, dass sie versucht „Faktoren des Erfolgs von Unternehmen zu identifizieren“ (Nicolai & Kieser, 2002, 579). Diese Definition entspricht im Kern den oben diskutierten Definitionen der Strategischen Managementforschung, wie sie insbesondere von US-amerikanischen Autoren vorgenommen wurden. Auch die Literaturreview in Kapitel 4 dieser Arbeit wird zeigen, dass die Aufdeckung von Kausalitäten zwischen der Performance und verschiedenen unabhängigen Variablen das Herzstück der Strategischen Managementforschung bildet. Die Trennung zwischen Erfolgsfaktorenforschung und Strategischer Managementforschung ist, so sie von Nicolai und Kieser überhaupt intendiert war, aus meiner Sicht daher künstlich und weitgehend irrelevant. Hiervon zu unterscheiden ist wiederum die Best Practice Forschung, die in der Tat auf die weitgehend unreflektierte Übertragbarkeit von Kausalitäten auf Settings, in denen diese Kausalitäten nicht nachgewiesen wurden, abzielen kann, wenn auch nicht muss; vergleiche beispielsweise Breene (2003), Buytendijk, (2006) oder Joyce, Nohria & Roberson (2004). Der Unterschied zwischen Best Practice-Forschung auf der einen und Strategischer Managementforschung auf der anderen Seite liegt im Fehlen beziehungsweise Vorhandensein aus Sicht einer deduktiv vorgehenden Wissenschaft belastbarer theoretischer Konstrukte zur Einordnung der generierten Erkenntnisse. Unter www.dialog-erfolgsfaktorenforschung.de findet sich der gesamte

nichtdeterministischen Beziehungen in der Managementforschung nicht gerecht werden können, Wissenschaftlichkeit zu signalisieren versucht, die tatsächlich kaum gegeben ist (Goshal, 2005).

Diese grundlegende Kritik an der Strategischen Managementforschung lässt sich erkenntnistheoretisch sowohl mit Poppers Argumentation zum kritischen Rationalismus als auch mit den instrumentalistischen Ansätzen Friedmans zumindest eingrenzen. Allgemein ist schwer nachzuvollziehen, warum ein Forschungsgebiet aufgegeben werden sollte, weil es bis dato (noch) keine voll befriedigenden Theorien und Methoden entwickelt hat, seine Forschungsfragen zu lösen. Klar ist, dass die Strategische Managementforschung aufgrund des komplexen und dynamischen Charakters des Gebietes wohl immer daran scheitern wird, die Leistung von Unternehmen universell zu erklären. Allerdings ist die Hypothese, dass es nicht doch aufdeckbare Kausalitäten geben kann, die die Performance von Organisationen zumindest in Teilen erklären, auch in keiner mir bekannten Weise falsifiziert worden. Daher ist ein Wissenszuwachs im Fachgebiet, nämlich ein besseres Verständnis des Kausalgeflechtes um das Performance-Konstrukt, nicht auszuschließen, wenn man grundsätzlich davon ausgeht, dass die Leistung einer Organisation nicht das Ergebnis eines Zufallsprozesses, also ausschließlich auf windfall profits zurückzuführen ist (vergleiche beispielsweise Schöler, 2004).

Zu beobachten ist allerdings das Auftreten eines konkurrierenden Paradigmas, das die zentrale Schwierigkeit der Strategischen Managementforschung, den Umgang mit Komplexität, grundlegend anders angeht. Aus Sicht der Organisationskybernetik und des System Dynamics Ansatzes (vgl. etwa Espejo, 1994, Pruckner, 2002, Schwaninger, 2004) ist bereits die grundlegende Annahme, dass komplexe Kausalitäten, wie sie dem Erfolg von Unternehmen zu Grunde liegen und in der Strategischen Managementforschung aufgedeckt werden sollen, über einfache kausale Modelle abbildbar sein könnten, zu verwerfen. Die Strategische Managementforschung angelsächsischer Prägung, wie sie im Rahmen dieser Arbeit untersucht wird und auch für den derzeitigen Diskurs im Fachgebiet verständnisleitend ist, basiert im Kern darauf, die Komplexität des Leistungserstellungsprozesses auf handhabbare Beziehungen zu reduzieren, in dem Bewusstsein, dass ein derartiger Ansatz notwendigerweise niemals der vollen Komplexität einer Unternehmung und ihrer Umwelt gerecht werden kann. Sie ist im Friedmanschen Sinne instrumentalistisch, versucht also nicht, die Realität abzubilden, sondern durch Komplexitätsreduktion „nützliche“ Vorhersagen zu generieren. Erkenntnis wird also durch Vereinfachung gewonnen. Die Organisationskybernetik³⁶ dagegen argumentiert, dass komplexe Systeme nicht verstanden werden können, wenn ihre Komplexität beschnitten wird. Der Wissenschaftler muss diesem Verständnis folgend also von Vereinfachungen absehen, um Erkenntnisfortschritt zu erreichen. Ein besonderes Augenmerk legt die Organisationskybernetik hierbei auf Interaktionsprozesse zwischen Individuen.

Beide Ansätze können in ihrem Umgang mit Komplexität, der zentral für die verwendeten Methoden in der Forschungspraxis ist, damit als sich gegenseitig ausschließende Paradigmen im Kuhn'schen Sinne verstanden werden (Kuhn, 2007). Wie Kuhn schreibt ist ein derartiger Paradigmenstreit kaum auflösbar, widerstreitende Paradigmen sind untereinander inkom-

Diskussionsstrang, der um Nicolai und Kiesers Artikel in der Zeitschrift „Die Betriebswirtschaft“ geführt wurde.

³⁶ Zum Zusammenhang zwischen Organisationskybernetik und der Theorie der Systemdynamik vergleiche Schwaninger, 2004.

messurabel. Inkommensurabilität bedeutet, dass Begriffe, Sätze und Theorien nicht über Paradigmengrenzen hinweg übertragbar sind (Kuhn, 2007). Entsprechend besteht für wissenschaftliche Arbeiten die Notwendigkeit, sich für die Begriffe und Zugänge eines Paradigmas zu entscheiden. Die Tatsache, dass die Strategische Managementforschung in dieser Arbeit verständnisleitenden Ansatz der Komplexitätsreduktion trotz all ihrer Probleme in der Lage ist, einen relevanten Anteil der Varianz ihrer zentralen abhängigen Variablen zu erklären deutet darauf hin, dass dieser Ansatz in seinem komplexitätsreduzierenden Verständnis eine Berechtigung hat und als Paradigma nicht verworfen werden muss. Insbesondere ist im Zuge des Paradigmenwettstreits auch bedeutsam, dass zentrale Publikationen im Fachgebiet bis dato nahezu ausschließlich dem klassischen Paradigma der Strategischen Managementforschung, dem der Komplexitätsreduktion, folgen. Im Zuge der in Kapitel 4 dieser Arbeit erstellten Literaturreview wurden keine empirischen Arbeiten augenfällig, die dem kybernetischen Paradigma folgen würden. Ein Grund hierfür kann auch darin liegen, dass es sich bei der Organisationskybernetik um ein sehr junges Forschungsfeld handelt und Instrumente und Methoden zur empirischen Überprüfung von Hypothesen in diesem Verständnis noch nicht befriedigend entwickelt sind. Ein Hinweis darauf sind die nach wie vor sehr geringen Zitationsraten, die Artikel aus der kybernetischen Tradition im Vergleich zu bedeutenden Beiträgen aus dem klassischen Verständnis der Strategischen Managementforschung aufweisen.

Festzuhalten bleibt in jedem Fall, dass das klassische instrumentalistische Paradigma der Strategischen Managementforschung, das mit Komplexitätsreduktionen Erkenntnis zu erlangen sucht, derzeit dominant und handlungs- und verständnisleitend ist. Die Heterogenität der Forschungsergebnisse der Strategischen Managementforschung, die das zentrale Erkenntnisinteresse der vorliegenden Arbeit darstellt, bezieht sich auf Forschungsergebnisse in der Tradition der klassischen Strategischen Managementforschung. Entsprechend konzentriert sich die vorliegende Arbeit im Folgenden auf die komplexitätsreduzierende Perspektive der Strategischen Managementforschung, ohne dass hiermit eine normative Aussage über den System Dynamics-Ansatz verbunden sein soll. In einer System Dynamics-Perspektive würden die in dieser Arbeit zu untersuchenden Problemfelder so nicht auftreten, da unbeobachtete Moderatorvariablen, die in Abschnitt 5 dieser Arbeit als wesentliche Treiber der Heterogenität im Forschungsgebiet verstanden werden, in einer systemdynamischen Perspektive von Anfang an in die entsprechenden Forschungen einbezogen werden müssten. Der System Dynamics Ansatz ist meines Wissens nach bisher nicht in größerem Umfang in der Strategischen Managementforschung zum Einsatz gekommen, entsprechend sind Aussagen über seinen empirischen Bewährungsgrad derzeit kaum möglich. Der den Ansatz charakterisierende Umgang mit hoher Komplexität stellt eine Hürde für empirische Arbeiten im kybernetischen Verständnis dar.

Fraglich bleibt, inwieweit mit dem Verständnis der klassischen Strategischen Managementforschung tatsächlicher Erkenntnisfortschritt gelingen kann. Analog der in Abschnitt 2.1 dargelegten Argumentation zur Möglichkeit des Zugewinnes von Wissen in den Erfahrungswissenschaften kann argumentiert werden, dass eine Verbesserung des Kenntnisstandes mit zunehmender Falsifikation überholter Theorien und einer stetigen Weiterentwicklung des Theoriekonglomerates durchaus denkbar und sinnvoll erscheint, solange nicht klar ist, dass der Weg der Quasiinduktion (in der Popperschen Terminologie) zu keinem Ziel führt. Dies ist, wie oben dargelegt, aus meiner Sicht nicht der Fall. Aus einer instrumentalistischen Sichtweise sollte auf dem vorhandenen Pfad so lange weitergearbeitet werden, bis bessere Alternati-

ven vorliegen. Diese könnten im kybernetischen Paradigma zu finden sein, allerdings ist dies bisher nicht absehbar. Die Suche nach, beziehungsweise das Ringen um diese besseren Alternativen ist der wesentliche Bestandteil der Managementforschung (vergleiche zum Instrumentalismus in der Strategischen Managementforschung auch Powell, 2002). Eine klare Ab-sage lässt sich aus erkenntnistheoretischer Sicht, analog der Argumentation in Abschnitt 2.1.1 und durchaus auch im Sinne der Kritik von Nicolai & Kieser (2002), allerdings für induktive Verfahren als Weg des Erkenntnisfortschritts ableiten. Die Komplexität und Heterogenität der Kausalbeziehungen lässt die Ableitung induktiver Schlüsse ohne Vorstrukturierung durch konsistente Axiomensysteme und Sätze schwerlich zu, die grundsätzliche Skepsis gegenüber induktiven Verfahren zur Wissensgenerierung gilt für das Fachgebiet der Strategischen Managementforschung aus meiner Sicht daher im Besonderen. Dies scheint auch breiter Konsens zu sein, wie der Blick in die Forschungspraxis zeigt. Während in qualitativ hochwertigen Journals rein deduktiv gearbeitet wird (vergleiche die in Kapitel 4 beschriebene Literaturreview), finden sich induktive Verfahren insbesondere im populären Bereich, etwa im theoretisch nicht weiter reflektierten Vergleich erfolgreicher Unternehmen, der viel geschmähten „Best Practice“-Forschung (etwa Breene & Nunes, 2006; Buytendijk, 2006; Joyce, Nohria & Rober-son, 2004; Reid & Hubbel, 2005; Peters & Waterman, 1982).³⁷ Umgekehrt gilt natürlich auch hier, dass nicht ausgeschlossen werden kann, dass aus Best-Practice Ansätzen „wahre“ Er-gebnisse, das heißt Aussagen über tatsächlich für eine befriedigend große Grundgesamtheit stabile Kausalbeziehungen entstehen, doch ist diese Möglichkeit vor dem Hintergrund der oben diskutierten Kritik an der Strategischen Managementforschung, die wesentlich auf die Komplexität des Untersuchungsfeldes abstellt, als unwahrscheinlich einzuschätzen. Die Stra-tegische Managementforschung scheint gut beraten, mit methodischer Disziplin und dedukti-ven Verfahren vorzugehen.

2.4 Zwischenfazit

In den ersten beiden Kapiteln der Arbeit wurden die Grundlagen der Untersuchung diskutiert. Die Problemlage und die Forschungsfragen wurden im ersten Kapitel vorgestellt, das zweite Kapitel beleuchtete die Möglichkeit des empirischen Erkenntnisgewinns aus einer wissen-schaftstheoretischen Perspektive und diskutierte und definierte den Terminus der Leistung, grenzte das Themenfeld der Strategischen Managementforschung ab und ordnete grundsätz-liche Kritiken am Fachgebiet ein.

Wissenschaftlicher Erkenntnisgewinn in den Sozialwissenschaften gelingt zuvorderst deduk-tiv, also theoriegeleitet. Grundgelegt ist die Präferenz deduktiven gegenüber induktiven Vor-gehens im Hume'schen Induktionsproblem, das logisch nachweist, dass aus Beobachtungen grundsätzlich keine Theorien (allgemeingültige Aussagen über die Realität) verifiziert werden können, da es immer weitere Theorien geben kann, die die Realität umfassender erklären. Dem setzt Popper das Postulat der Falsifikation entgegen. Erkenntnisgewinn entsteht dem-nach durch die Falsifikation von Theorien, die so durch stetig bessere ersetzt werden. Not-wendig ist die Generierung von Theorien und deren Test an der Realität, also ihre empirische

³⁷ Die Möglichkeit, induktive Verfahren für explorative Ideenfindungsprozesse einzusetzen, bleibt hier-von natürlich unberührt. Vergleiche zur Best Practice-Forschung im Fachbereich Strategisches Mana-gement auch Bromiley & Johnson (2005) und Simon (1946).

Überprüfung. Hält eine Theorie dieser Überprüfung nicht stand, ist sie zu verwerfen. Die kritisch-rationalistische Sicht Poppers ist in der Strategischen Managementforschung sicher verständnisleitend, wiewohl in der wissenschaftlichen Praxis ein deutlich instrumentalistischeres Verständnis von Theorien im Sinne Friedmans vorherrscht. Demnach sind Theorien so lange beizubehalten, wie sie sinnvolle Beiträge zur Erklärung der Realität liefern, auch wenn sie auf falsifizierten oder von vorneherein unrealistischen Axiomen basieren. Beispielhaft für diese Auseinandersetzung ist das Ringen um und zwischen den derzeit dominanten Theorien der Strategischen Managementforschung, dem Resource- und Market-based view. Klar ist jedoch, dass Popper Skepsis gegenüber rein induktiven Verfahren geteilt wird, eine theoretische Fundierung als Ausgangspunkt jedweden empirischen Erkenntnisstrebens ist in der aktuellen Forschungslandschaft unumgänglich. Der Umgang mit falsifizierten Theorien entspricht jedoch mehr den Empfehlungen Friedmans als denen Poppers, sie werden so lange beibehalten, wie sie sinnvolle Beiträge zum Verständnis der komplexen Kausalitäten in der Strategischen Managementforschung leisten. Nichtsdestotrotz kommt dem Hypothesentest als wesentlichem Instrument der deduktiven Wissenschaft, also der Herabführung allgemeiner in besondere Sätze, eine Schlüsselrolle im Prozess der Wissensgenerierung in der Strategischen Managementforschung zu. Zur Aufdeckung kausaler Beziehungen ist die präzise Spezifikation von Subjekt und Objekt der Wenn-Dann-Sätze entscheidend. Die Leistung einer Organisation ist die dominante abhängige Variable in der Strategischen Managementforschung, ihre Spezifikation durch den Prozess der Operationalisierung ist damit von höchster Wichtigkeit für Erkenntnisfortschritte im Fachgebiet. Der Verdacht, dass eine Fehlspezifikation dieser zentralen Variablen für einen Teil der beobachteten Unfähigkeit des Fachgebietes, reproduzierbare Kausalitäten aufzuzeigen, verantwortlich sein könnte, bleibt damit weiterhin im Raum.

Der Leistungs- und im englischen Sprachraum auch der Performancebegriff sind von einer erheblichen Bedeutungsunschärfe umgeben, die behoben werden muss, um eine befriedigende Konstruktvalidität zu erreichen. Im Rahmen dieser Arbeit wird Leistung definiert als die Summe des Ergebnisses der Tätigkeiten einer Organisation. Mit dieser Definition wird insbesondere ein tätigkeitsorientiertes Leistungsverständnis ausgeschlossen. Das ergebnisorientierte Leistungsverständnis ist die Voraussetzung für alle weiteren Betrachtungen und wird implizit in allen Arbeiten zur Strategischen Managementforschung, die im Rahmen der Literaturreview Beachtung fanden, vorausgesetzt. Von Bedeutung ist auch die Abgrenzung zum Business Performance Measurement, dem leistungsbezogenen Messinstrumentarium zur Unternehmenssteuerung. Im Rahmen der Strategischen Managementforschung verbieten die spezifischen Charakteristika von Hypothesentests in diesem Feld die Anwendung der meisten Instrumente des Business Performance Measurement. Es ist notwendig, sich auf Indikatoren zu beschränken, die für eine relativ große Zahl an Untersuchungsobjekten relativ einfach erfassbar sind.

Schlussendlich erfolgte im zweiten Kapitel der Arbeit noch eine Abgrenzung des Fachgebietes der Strategischen Managementforschung. Gegenstand dieses Forschungsfeldes ist die Erklärung des Verhaltens von Unternehmen auf der strategischen Ebene, die Erklärung von Unterschieden in der Leistungsfähigkeit sowie die Erarbeitung normativer Aussagen aus den gewonnenen Ergebnissen. Von einigen Autoren wird eine grundlegende Skepsis dahingehend geäußert, ob es prinzipiell möglich sein kann, die Forschungsziele des Fachgebietes zu errei-

chen. Grund dieser Skepsis ist die hohe Komplexität sowie die Historizität des Untersuchungsfeldes. Diese Skepsis mag berechtigt sein, doch scheint die Argumentation nicht derart stichhaltig, dass sie eine Preisgabe des Fachgebietes erforderlich machen würde. Vielmehr ist aus den bisher erarbeiteten erkenntnistheoretischen Betrachtungen heraus eine methodisch saubere deduktive Vorgehensweise geboten, mit der dann, trotz aller fachimmanenten Schwierigkeiten, auch ein Wissenszuwachs möglich sein sollte. Mit dem kybernetischen Ansatz steht ein konkurrierendes Paradigma bereits bereit, das eine zentrale Schwierigkeit der klassischen Strategischen Managementforschung, den Umgang mit Komplexität, grundsätzlich anders angeht. Welches der beiden Paradigmen sich durchsetzen wird, wird der weitere Wettstreit der Theorien zeigen. In der empirischen Verwendung liegen die Vorteile hier derzeit klar beim klassischen Ansatz.

Im Folgenden gilt es nun, verschiedene Erklärungsansätze für die beobachtete Heterogenität der Ergebnisse in der Strategischen Managementforschung konzeptionell zu beleuchten. Drei Problemfelder sind hierbei von besonderem Interesse: Zum ersten das herausgearbeitete Problem der Leistungsoperationalisierung. Hierbei ist konzeptionell zu erschließen, aufgrund welcher Systematik ein Zusammenhang zwischen Operationalisierung und Studienergebnis bestehen könnte. Zum zweiten ist das Problem der begrenzten Aussagemacht von Studien zu untersuchen. Hier steht die Vermutung im Raum, dass sich Arbeiten zwar mit denselben Zusammenhängen beschäftigen, diese aber in unterschiedlichen Teilpopulationen der Grundgesamtheit aller Unternehmen beleuchten und durch die Heterogenität zwischen diesen Teilpopulationen Heterogenität in den Forschungsergebnissen entsteht. Zum Dritten ist das Problem unklarer zeitlicher Zusammenhänge zwischen Ursache und Wirkung zu beleuchten. Zwischen Ursache und Folge vergeht Zeit, dies wird in Forschungsarbeiten möglicherweise nicht ausreichend berücksichtigt.

3. Theoriegeleitete Analyse

3.1 Überblick

Für echten Erkenntnisfortschritt wäre es ideal, die Gesamtheit aller Unternehmen in ihrer gesamten Tiefe zu jeder beliebigen Zeit betrachten zu können (Carroll, 1994). Aufgrund der Vielzahl der Akteure, des dynamischen und damit historischen Charakters des Wirtschaftsprozesses und der Komplexität der Vernetzung der Akteure ist dies naturgemäß ein nicht erfüllbarer Wunsch. Der Wissenschaftler ist so gezwungen, immer Ausschnitte der Wirklichkeit zu untersuchen. Zugleich muss er sich bewusst sein, dass er selbst in diesen Ausschnitten nur Teile der Vernetzungen und Kausalitäten analysieren kann. Je mehr Teile und Kausalitäten untersucht werden sollen oder je breiter der Wirklichkeitsausschnitt der Untersuchung gewählt wird, desto aufwändiger wird das entsprechende Untersuchungsdesign (Combs et al., 2005). In diesem Zusammenhang ist ein notwendiger Trade-off zwischen Breite und Tiefe der Untersuchung zu bewältigen (Miller, 2005). In Konsequenz muss sich der Forscher also entscheiden, wo und wie er seine begrenzten Ressourcen einsetzen will, das heißt wie viel Anwendbarkeit er für einen gewissen Zugewinn an Validität aufzugeben bereit ist und umgekehrt. Mit dieser Einschränkung der Allgemeinheit geht jedoch ein Verallgemeinerungsproblem der gewonnenen Aussagen einher, das im folgenden Abschnitt 3.2 diskutiert wird (vergleiche hierzu beispielsweise auch Hall & Tickle-Degnen, 1994).

Wie in den Abschnitten 2.1 und 2.2 erarbeitet stellt die Operationalisierung komplexer Konstrukte eine zentrale Herausforderung empirischer Forschungen dar. Zu fordern ist die Verwendung von Indikatoren, die in der Lage sind, das Konstrukt, das sie messen sollen, valide abzubilden. Fraglich ist jedoch, ob dieser Forderung in der Strategischen Managementforschung im Rahmen ihrer wichtigsten Operationalisierung, der Messung der Leistung von Organisationen, ausreichend nachgekommen wird. Dies wird von verschiedenen Autoren bestritten (Starbuck, 2004; Combs et al., 2005). Ein wesentlicher Grund hierfür liegt, wie in Abschnitt 2.2.1 diskutiert, in der Diffusität beziehungsweise Mehrdimensionalität des Leistungsbegriffs, der alles vom schieren Überleben des Unternehmens über finanzielle bis hin zu Stakeholder-orientierten Kennzahlen umfassen kann.³⁸ Die Forderung nach konvergenter Validität würde eine hohe Korrelation zwischen diesen Operationalisierungen erwarten lassen. Wie gezeigt werden konnte (Rowe & Morrow Jr., 1999) ist dies nicht notwendigerweise der Fall, vielmehr lassen sich Performanceoperationalisierungen auf unterschiedliche Art und Weise clustern, sie weisen also eine gewisse Diskriminanz auf, die Forderung nach konvergenter Validität ist regelmäßig verletzt.³⁹ Ein Grund hierfür könnte im Prozess der Verteilung der Wertschöpfung einer Organisation zwischen Stakeholdergruppen und den differierenden Punkten an diesem Prozess, an denen die verschiedenen Operationalisierungen der Leistung messen, zu finden sein. Wäre dem so, würde die Forderung nach konvergenter Validität erreicht werden, indem Leistungsoperationalisierungen gewählt werden, die die gleiche Stelle dieses Appropriationsprozesses erfassen. Umgekehrt würde eine Leistungsoperationalisierung, die unreflektiert eine beliebige Stelle herausgreift, nicht nur die Leistung einer Organi-

³⁸ Das Spektrum tatsächlich verwendeter Leistungsoperationalisierungen ist Anhang A.2 zu entnehmen.

³⁹ Diese Clusterung basiert auf einer statistisch signifikanten Faktorenanalyse, vergleiche Rowe & Morrow Jr. (1999).

sation messen sondern gegebenenfalls zudem die Appropriationsfähigkeit einer bestimmten Stakeholder-Gruppe. Die Prozesse der Leistungserstellung und der Leistungsverteilung wären konzeptionell zu trennen. Damit könnte sich das Phänomen erklären lassen, dass unterschiedliche Leistungsoperationalisierungen unterschiedliche Kausalitäten anzeigen, ein zentrales Erkenntnisinteresse dieser Arbeit (vergleiche hierzu Kapitel 1). Die relative Verhandlungsmacht der Stakeholder-Gruppen würde damit eine latente Variable darstellen, die kausale Aussagen im Zusammenhang mit der Leistung von Organisationen beeinflussen kann. Dieser Argumentationsstrang wird in Abschnitt 3.3 vertieft.

Kausale Verbindungen zwischen abhängigen und erklärenden Variablen sind im Kontext der Strategischen Managementforschung auch deshalb schwer aufzudecken, weil der zeitliche Zusammenhang zwischen Ursache und Wirkung nicht immer unmittelbar ersichtlich ist. Wenn unklar ist, ob ein Ereignis X in Periode t_0 , t_1 oder t_2 zu einer Veränderung der abhängigen Variablen führt, kann eine zeitliche Fehlspezifikation des Zusammenhangs in einer statistischen Auswertung dazu führen, dass Kausalitäten übersehen werden (Bowen & Wiersema, 1999; Zapf, Dormann & Frese, 1996; Brammer & Millington, 2008). Im Extremfall kann dies auch zu Fehlschlüssen hinsichtlich der Richtung der Kausalität führen, also das Phänomen der reversed causalities auftreten, der umgekehrten Kausalitäten. Ist dies der Fall lässt sich zwar ein statistischer Zusammenhang zwischen abhängiger und erklärender Variablen finden, doch ist dieser, anders als die Ausdeutung des Zusammenhangs nach den üblichen statistischen Verfahren nahelegt, so, dass tatsächlich die abhängige Variable die erklärende beeinflusst und nicht umgekehrt.⁴⁰ Dieses Problem ist wohlbekannt und wird auch in der aktuellen Forschungspraxis wahrgenommen (Hmieleski & Baron, 2009; Miller, Fern & Cardinal, 2007; Morrow Jr., Sirmon, Hitt & Holcomb, 2007; Short et al., 2007; Sun, Aryee & Law, 2007; Hall & Tickle-Degnen, 1994), wenn auch selten systematisch diskutiert. Mit Längsschnittdatensätzen, Panelanalysen und pfadanalytischen Verfahren liegen eine Reihe Möglichkeiten vor, das Problem der unklaren zeitlichen Zusammenhänge zwischen Ursache und Wirkung in den Griff zu bekommen, diese werden jedoch in der Forschungspraxis nicht konsequent eingesetzt (vergleiche Bowen & Wiersema, 1999 und Abschnitt 4.2.4). Im Abschnitt 3.4 wird das Problem konzeptionell beleuchtet.⁴¹

3.2 Das Problem der Populationsselektion

3.2.1 Populationsselektion – Die Problemlage

Wie bereits in Abschnitt 3.1 diskutiert, würde maximale Validität der Erkenntnisse nur aus der Betrachtung der vollständigen Grundgesamtheit der Untersuchungsobjekte zu erwarten sein

⁴⁰ Genau genommen treffen übliche statistische Verfahren natürlich überhaupt keine Aussagen zur Richtung des aufgezeigten Zusammenhangs, in der Praxis wird jedoch in aller Regel davon ausgegangen, dass eine signifikante Beziehung zwischen abhängiger und erklärender Variablen eine Bestätigung der überprüften Hypothese darstellt, also die postulierte Beeinflussung der abhängigen durch die erklärende Variable und nicht umgekehrt vorliegt. Sharfman & Fernando (2008) und Chang & Xu (2008) diskutieren dieses Problem in ihren Beiträgen.

⁴¹ Diese drei im Hauptteil der Arbeit untersuchten Problemfelder sind zentral, aber nicht erschöpfend. Weitere in der Literatur diskutierte Methodenprobleme finden sich mit einer kurzen Einordnung in Anhang A.3.2.2. Hunter & Schmidt (2007) erstellen darüber hinaus eine umfassende Kategorisierung von Methodenproblemen im Rahmen quantitativer Erhebungen, hierzu finden sich Anmerkungen in Anhang A.3.2.1.

(Carroll, 1994; Short et al., 2002). Die Reduktion der Grundgesamtheit bewirkt in Konsequenz, dass der Wissenschaftler mit jeder Einschränkung von weniger Fällen ausgehend auf das Allgemeine schließt, eine Vorgehensweise, die sowohl aus Sicht eines erklärenden als auch eines verstehenden (oder hermeneutischen) Wissenschaftsverständnisses abzulehnen ist (Wolf, 2008, 17f; vergleiche auch Abschnitt 2.1.1). Diese Einschränkung der Allgemeinheit geschieht zumeist bewusst durch die Eingrenzung der Population, wenn beispielsweise die Untersuchung auf eine bestimmte Branche fokussiert wird, um unbeobachtete Brancheneffekte auszuschließen und so die Homogenität des resultierenden Samples und die Erklärungsmacht der statistischen Modelle zu erhöhen, teilweise aber auch unbewusst oder zumindest nicht expliziert. Dies ist etwa der Fall, wenn eine Variablenspezifikation eine Reduzierung der Grundgesamtheit bedingt, ohne dass diese Reduzierung intendiert oder erwünscht ist. Eine solche systematische implizite und unerwünschte Reduktion der Grundgesamtheit tritt beispielsweise bei marktbasierenden Leistungsindikatoren auf. Marktbasierende Kennzahlen setzen naturgemäß eine öffentliche Notierung der Anteilsscheine der zu untersuchenden Unternehmen an Kapitalmärkten voraus, sowie eine ausreichend hohe Marktkapitalisierung, um die Herausbildung eines Marktpreises zu ermöglichen.⁴²

An der Deutschen Börse waren zum 30. Juni 2009 1190 Unternehmen notiert.⁴³ Demgegenüber erfasste das Statistische Bundesamt für das Jahr 2006 gut 3,5 Mio. gemeldeter Unternehmen in Deutschland.⁴⁴ Leistungsoperationalisierungen, die auf börsenwertbasierten Kennzahlen beruhen, schließen damit rechnerisch 99,97 Prozent aller Unternehmen in Deutschland⁴⁵ aus. Selbst unter Nichtberücksichtigung der 2,3 Mio. Einzelunternehmer verbleibt noch eine Quote von 99,91 Prozent. Reduziert man die Betrachtung auf GmbHs und AGs, schließt also Kleinstunternehmen, insbesondere Personengesellschaften, aus, liegt der Anteil der nicht börsennotierten immer noch bei 99,7 Prozent. Auch wenn neben der Deutschen Börse weitere Börsenbetreiber in Deutschland existieren ist festzuhalten, dass die allermeisten Unternehmen bei Verwendung marktbasierter Kennzahlen systematisch aus der jeweiligen Untersuchung ausgeschlossen werden, ohne dass diese Exklusion intendiert ist.⁴⁶

⁴² Ein Sonderfall bei den marktbasierenden Kennzahlen sind die Kapitalkosten. Hier wird nicht der Marktwert des Unternehmens durch die Kapitalmärkte bewertet, sondern seine Bonität, also in Konsequenz die Einschätzung der Marktteilnehmer über die Zahlungs- beziehungsweise Überlebensfähigkeit des zu untersuchenden Unternehmens. Hierfür ist nicht notwendigerweise eine öffentliche Notierung der Anteilsscheine des Unternehmens erforderlich. Dieser Indikator findet jedoch in der Strategischen Managementforschung nur sehr selten Anwendung (vergleiche Kapitel 4).

⁴³ Prime-, General- und Entry-Standard sowie First Quotation Board. Nicht in dieser Zählung enthalten sind ausländische Unternehmen mit Zweitlisting an der Deutschen Börse, also Anteilsscheine im Second Quotation Board. Quelle: www.deutsche-boerse.com.

⁴⁴ Quelle: Statistisches Bundesamt (www.destatis.de), Statistik: Unternehmensregister.

⁴⁵ Die Zahlen für die USA, der für die Strategische Managementforschung relevantesten Volkswirtschaft (vergleiche Anhang A.2.2): An der weltgrößten Börse NYSE/Euronext sind 4.100 Unternehmen gelistet, in den USA 5,9 Mio. Unternehmen gemeldet, der Anteil der an der NYSE notierten Unternehmen liegt bei 0,07 Prozent und damit etwas höher als in Deutschland. Quellen: www.nyse.com und U. S. Census Bureau, 2007. Listing-Zahlen von 2008, Unternehmensmeldungen von 2005. Wie im Folgenden für die Situation in Deutschland diskutiert ist diese Zahl unscharf, durch die schiere Größenordnung jedoch relevant.

⁴⁶ In relativen Zahlen reduziert sich dieses Problem natürlich, da börsennotierte Unternehmen ökonomisch aufgrund ihrer zumeist erheblichen Größe von größerer Relevanz sind. So liegt etwa der Anteil der Beschäftigten in börsennotierten Unternehmen an der Gesamtbeschäftigung deutlich höher als dies der Anteil der börsennotierten Unternehmen an den Gesamtunternehmen suggerieren mag. Die grundsätzliche Beobachtung, dass börsennotierte Unternehmen nur einen geringen Ausschnitt der

Hiermit geht naturgemäß ein Verlust an Verallgemeinerbarkeit, an Erklärungsmacht (Popper, 1989), einher. Marktbasierte Kennzahlen stellen in diesem Sinne den Extremfall dar, aber auch viele andere Kennzahlen können eine ähnliche Wirkung entfalten, und sei es nur aus forschungspraktischen Gründen. So werden etwa kleine und mittlere Unternehmen regelmäßig ihre Bilanzen nicht öffentlich zugänglich machen. Dies begrenzt wiederum die Möglichkeit der Berechnung von Kapitalrentabilitäten, der mit Abstand am häufigsten angewandten Operationalisierung von Leistung (vergleiche 4.2.3), auf größere Unternehmen. Damit ist aufgrund der Ergebnisse der qualitativen Literaturreview ein erheblicher Bias der Strategischen Managementforschung in Richtung großer Unternehmen anzunehmen.

Problematisch ist hierbei weniger die Existenz dieser Reduktion der Gesamtpopulation auf eine kleinere Teilpopulation als vielmehr die Suggestion einer nicht begründeten Aussagemacht der getesteten Hypothesen, die zu Fehlfolgerungen sowohl für die normativ-praktische Dimension der Wirtschaftswissenschaften als auch und insbesondere für die Fortführung des wissenschaftlichen Diskurses führen kann (Hall & Tickle-Degnen, 1994). Zu fordern ist, dass die Aussagemacht der Hypothesentests und ihre Grenzen von den Autoren wissenschaftlicher Beiträge expliziert werden. Wie die Auswertung der qualitativen Literaturreview in Kapitel 4 zeigen wird, ist dies jedoch bei weitem nicht immer der Fall.

Die Einschränkung der Aussagemacht ist keineswegs grundsätzlich kritisch zu sehen. Die Reduktion der Grundgesamtheit ist aufgrund der nicht beherrschbaren Größe der Gesamtpopulation eine zwingende Notwendigkeit (vergleiche auch Wan & Yiu, 2009). Zudem ist die Einschränkung der Population insbesondere dann unproblematisch, wenn sich die erklärenden Variablen auf den gleichen Anteil der Grundgesamtheit beschränken wie die abhängige. Um das Beispiel der Börsennotierung fortzuführen läge dann keine schädliche Einschränkung der Verallgemeinerbarkeit bei der Verwendung marktbasierter Kennzahlen vor, wenn auch die erklärenden Variablen marktbasierend sind, wenn etwa die Verbreitung von Aktienoptionsprogrammen untersucht werden soll, die ihre volle Motivationswirkung nur bei öffentlich notierten Unternehmen entfalten (Sanders & Hambrick, 2007); wohl aber kann die Begrenzung problematisch sein, wenn beispielsweise das Risikoverhalten von CEOs im Fokus der Betrachtung steht (Shimizu, 2007). In letzterem Fall ist die intendierte Absicht nicht, lediglich eine Aussage über CEOs börsennotierter Unternehmen zu machen sondern allgemein das Verhalten von Geschäftsführern zu verstehen, in ersterem Fall geht es tatsächlich lediglich um börsennotierte Organisationen. Die Einschränkung der untersuchten Unternehmen auf eine bestimmte Größenklasse, um ein zweites Beispiel anzuführen, ist dann keine Eingrenzung der Aussagemacht der zu untersuchenden Hypothese, wenn beispielsweise in Hmieleski & Baron (2009) spezifisch junge Unternehmen untersucht werden sollen, wohl aber, wenn McDonald, Khanna & Westphal (2008) den Einfluss externer persönlicher Netzwerke von CEOs mit der Performance ihrer Unternehmen in Beziehung setzen und sich hierbei auf Großunternehmen beschränken. In ersterem Fall ist die Einschränkung auf junge Unternehmen ein Bestandteil der Forschungsfrage, in letzterem eine nicht intendierte Einschränkung der Allgemeingültigkeit der Aussage.

ökonomischen Wirklichkeit darstellen, bleibt hiervon jedoch unberührt, ebenso die Vermutung, dass diese systematische Exklusion bei bestimmten Kennzahlen für einen Teil der Heterogenität von Studienergebnissen in der Strategischen Managementforschung verantwortlich sein könnte.

Das Qualitätskriterium kann also nicht die Existenz oder Vermeidung von Selektionsfiltern oder die Größe und der Umfang der resultierenden Teilpopulation sein (obwohl eine breite und tiefe Teilpopulation *ceteris paribus* natürlich zu präferieren ist), sondern die Explikation der Aussagemacht der geprüften Hypothesen. Heterogene Teilpopulationen weisen zunehmend mehr unerklärte Einflüsse auf, je größer die Zahl der Dimensionen wird, in der sich einzelne Subjekte in der Population voneinander unterscheiden. Die Aufdeckung kausaler Aussagen mit Hilfe statistischer Methoden wird damit immer schwieriger (vergleiche Bowen & Wiersema, 1999). Umfasst eine in einer Studie zu untersuchende Teilpopulation etwa Unternehmen unterschiedlicher Größen und Branchen, die aus Wirtschaftsräumen mit differierenden Regulierungsrahmen und Wertvorstellungen stammen, werden kausale Zusammenhänge kaum mehr auffindbar sein. Eine Eingrenzung der Population ist damit immer eine forschungspraktische Notwendigkeit und muss je nach zu untersuchender Forschungsfrage individuell beleuchtet werden. Diese Einschränkung der Grundgesamtheit kann jedoch durchaus auch die Heterogenität von Forschungsergebnissen in der Strategischen Managementforschung erklären. Im Rahmen der Meta-Analyse in Kapitel 5 wird zu überprüfen sein, inwieweit dies tatsächlich der Fall ist.⁴⁷

Warum ist die Einschränkung der Grundgesamtheit kritisch? Diese Frage soll im Folgenden detaillierter betrachtet werden. Im Rahmen der Auswertung der qualitativen Literaturreview gibt Abschnitt 4.2.3 dann einen Überblick über in der aktuellen Strategischen Managementforschung verwendete Selektionsfilter⁴⁸ und resultierende Teilpopulationen sowie über die Häufigkeit, mit der die hieraus resultierende begrenzte Verallgemeinerbarkeit gewonnener Studienergebnisse diskutiert wird.

3.2.2 Populationsselektion – wissenschaftstheoretisches Problem

Die Rolle der Empirie in der Wirtschaftswissenschaft wurde bereits in Abschnitt 2.1 ausführlich diskutiert. Grundsätzlich gefährdet das oben angerissene Problem der Populationsselektion das Allgemeingültigkeitspostulat der mit den gezogenen Stichproben zu prüfenden Theorien. Das heißt, dass die empirische Studie nicht, wie für einen Theorietest aus kritisch-rationalistischer Sicht zu fordern, überprüfen kann, inwieweit das Aussagesystem der Theorie die Realität abbilden kann, sondern lediglich einen Sonderfall der Allgemeinheit überprüft. Je spezieller dieser Sonderfall, desto weniger wertvoll ist die verbundene Empirie für den Test einer Theorie, oder, um mit Popper (1989) zu sprechen, desto weniger kann sich die zu Grunde liegende Theorie bewähren. Ganz konkret kann die (bewusste oder unbewusste) Einschränkung der Grundgesamtheit im erkenntnistheoretischen Erklärungszusammenhang als Schwächung der Erklärungsmacht der aus einem Satz abgeleiteten zu prüfenden Hypothese verstanden werden. Eine gehaltvolle Hypothese, ein synthetisches Urteil in der Kant'schen Dichotomie (vgl. Abschnitt 2.1; Kant & Eisler, 1930/2009; Maurer, 2004), besteht (wenigstens) aus Subjekt und Objekt, die mittels einer Wenn-Dann-Beziehung verknüpft sind. Die

⁴⁷ Vergleiche zur Abgrenzung des Populationsselektionsproblems zum etwa im Standardwerk von Hunter & Schmidt (2007) diskutierten Problem des Sampling Errors die Ausführungen in Anhang A.3.2.2.

⁴⁸ Das rein technische Problem der Stichprobenziehung wird im Rahmen dieser Arbeit nicht weiter beleuchtet, der interessierte Leser sei auf Lamnek (1995) in Kombination mit Hartung et al. (2009) verwiesen.

Hypothese ist umso gehaltvoller, je umfassender die Antezedens (der „wenn-Satz“) und je spezifischer die Konsequenz (der „dann-Satz“) ausfallen.

An einem populären Beispiel lässt sich das Problem veranschaulichen: In einer der klassischsten Untersuchung zur Leistungsfähigkeit von Unternehmen, „In Search of Excellence“ (Peters & Waterman, 1982), formulieren die Autoren unter anderem sinngemäß die Hypothese:

Ein Unternehmen wird dann zum Hochleister, wenn es eine Bindung an das Kerngeschäft beibehält.

Diese Aussage ist ein gehaltvoller synthetischer Satz in der Kant'schen Sprachregelung (Kant & Eisler, 1930/2009), die Aussage gilt für alle Unternehmen und ist entsprechend gut falsifizierbar, der Erklärungswert ist hoch. Nun jedoch betrachten die Autoren ausschließlich Unternehmen aus einem Sample von 62 Organisationen, die von Experten als besonders erfolgreich eingeschätzt wurden. Hierbei handelt es sich ausschließlich um Konzerne erheblicher Größe („the usual suspects“, Katzenbach, 2000), ausgeblendet bleiben kleinere, weniger bekannte Unternehmen, die wesentliche Mehrzahl der Objekte aus der Grundgesamtheit aller Unternehmen. Es handelt sich um börsennotierte US-amerikanische Unternehmen. Diese Eingrenzung wirkt damit als Einschränkung der Antezedens; die tatsächlich getestete Hypothese lautet also:

US-amerikanische börsennotierte Großunternehmen, die sich bereits einen guten Ruf erarbeitet haben, werden dann zu Hochleistern, wenn sie eine Bindung an das Kerngeschäft beibehalten.

Die Hypothese wird damit entwertet, gilt nun nicht mehr für eine Vielzahl an Unternehmen, sondern nur mehr für einen, auf die Gesamtpopulation der Unternehmen bezogen, sehr geringen Anteil. Die Erklärungsmacht und damit der wissenschaftliche und praktische Wert der Hypothese und der mit ihr verbundenen Aussagen sinken. Dass es sich bei diesen Überlegungen nicht um ein leeres Gedankenspiel handelt belegt etwa die Studie von Chakrabarti, Singh & Mahmood (2007). Die Autoren untersuchen hier die oben diskutierte These, dass sich Diversifikation negativ auf die Performance auswirkt, in unterschiedlichen Volkswirtschaften. Sie finden einen negativen Zusammenhang zwischen Performance und Diversifikationsgrad in hoch entwickelten, jedoch einen positiven Zusammenhang in weniger entwickelten Volkswirtschaften. Zahlreiche weitere Beispiele lassen sich finden, der Umgang mit der Problemlage in der aktuellen Forschung wird im Rahmen der Literaturreview in Abschnitt 4.2.3 detaillierter ausgeführt. Aus einer wissenschaftstheoretischen Sichtweise stellt eine zu spezifisch gewählte Grundgesamtheit also die Möglichkeit in Frage, die Theorie zu bewähren, die getestet werden soll.

3.2.3 Populationsselektion – Literaturüberblick

Betrachtet man die relevantesten Kriterien, die einer Reduzierung der Grundgesamtheit hin zu kleineren Teilpopulationen zu Grunde liegen wird schnell klar, dass eine Verallgemeinerbarkeit von Ergebnissen in Teilpopulationen auf eine größere Gesamtpopulation regelmäßig nicht zu erwarten sein wird. Fünf Selektionskriterien sind in der Strategischen Managementforschung zentral: Die Branche, die geografische Herkunft der untersuchten Unternehmen, der Zeitraum der Erhebung, die Größe und der Notierungsstatus der einbezogenen Unternehmen sowie, ein weiterer nicht intendierter Selektor, Datenbanken, aus denen

die einer Studie zu Grunde liegenden Indikatorenwerte zumindest teilweise entnommen werden können. Alle fünf sollen im Folgenden knapp diskutiert werden.

Branche

Die Relevanz des industriellen Kontextes, der Branche, auf die Performance von Unternehmen ist allgemein anerkannt und wird etwa in Short et al. (2007) ausführlich diskutiert. In einer Studie, die Manager aus einem kognitionspsychologischen Kontext heraus in Strategie-Typen einteilt, finden beispielsweise auch Kabanoff & Brown (2008) einen starken Einfluss der betrachteten Branche auf die abhängige Variable. Auch Hawawini, Subramanian & Verdin (2003) betonen, dass sich für die meisten Unternehmen ein wesentlicher Teil ihrer Performance allein durch die Branche erklären lässt. Eine Ausnahme hiervon sind den Autoren zufolge ausschließlich einige wenige besonders leistungsfähige oder wertvernichtende Unternehmen in jeder Branche. Im Extremfall erklärt in dieser Studie von Hawawini und Kollegen die Branchenzugehörigkeit 30 Prozent der Performance-Unterschiede zwischen Unternehmen, während die Gesamtheit aller unternehmensspezifischen Faktoren lediglich 17 Prozent der Varianz der Leistung von Unternehmen erklärt. Einen Hinweis auf die Relevanz der Branche liefert beispielsweise auch Powell (2003b): Der Autor vergleicht die Performance-Unterschiede in 21 Industrien und identifiziert insbesondere die Pharmaziebranche, die Brauindustrie und die Computerbranche als von besonders ausgeprägten Performance-Differenzen zwischen in diesen Branchen agierenden Unternehmen gekennzeichnet, während sich in anderen Branchen ein relativ homogenes Performancebild ergibt. Diese besonders leistungsheterogenen Branchen wiederum werden sehr häufig herangezogen, wenn Performance-Differenzen generell erklärt werden sollen (Powell & Arregle, 2007). Auch Barney & Mackey (2005) weisen auf die Verallgemeinerungsprobleme hin, die entstehen, wenn lediglich kleine Populationen von Unternehmen in spezifischen Industrien betrachtet werden. Im Rahmen einer spezifischeren Performance-Betrachtung diskutieren Berrone & Gomez-Mejia (2009) den Zusammenhang zwischen CEO-Bezahlung und der Leistung im Bereich Vermeidung von schädlichen Emissionen. Die Autoren folgern aus ihren und den Ergebnissen vergleichbarer Studien, dass ein positiver Einfluss nur dann zu erwarten ist, wenn die untersuchten Unternehmen in Branchen agieren, die sich starkem öffentlichem Druck zur Emissionsvermeidung ausgesetzt sehen.

Geographische Herkunft

Die Bedeutung nationaler Unterschiede für Kausalbeziehungen in der Strategischen Managementforschung kann insbesondere aufgrund zahlreich vorliegender Erkenntnisse verwandter Disziplinen als erwiesen gelten.⁴⁹ Bedeutend ist sowohl die nationale Kultur, die etwa die

⁴⁹ Außerhalb des Gebietes der Strategischen Managementforschung befasst sich die volkswirtschaftlich ausgerichtete Forschung zu Regional Economics (beispielsweise Storper, 1995) mit den Auswirkungen von Region und Raum auf Organisationen. Studiert werden hier allerdings relativ kleinteilige Räume von der Größe deutscher Raumordnungsregionen (vergleiche auch Fritsch, Henning, Slavtchev & Steigenberger, 2007). Die Varieties of Capitalism-Literatur erforscht Unterschiede in den Ausprägungen wirtschaftlicher Grundordnungen und ist damit in diesem Kontext ebenfalls relevant (vergleiche beispielsweise Hall & Soskice, 2004). Das Fachgebiet der interkulturellen Kommunikation befasst sich ebenfalls mit Unterschieden zwischen Regionen, aus einer individuellen, kulturanthropologischen und

Bedeutung und Art persönlicher Beziehungen beeinflusst, als auch der ordnungspolitische Rahmen. Ersteren Fall, die kulturelle Dimension, untersucht beispielsweise Luo (2008a). Der Artikel beleuchtet die Rolle von Vertrauen in Netzwerkbeziehungen und wählt China als Untersuchungsfeld. Die Auswahl dieses Feldes wird unter anderem damit begründet, dass Vertrauen für chinesische Arbeitnehmer von besonders hoher Bedeutung ist. Diese Hypothese kann auch bestätigt werden. Zu einem ähnlichen Ergebnis kommen auch Chakrabarti et al. (2007) im Rahmen ihrer Untersuchung des Zusammenhangs zwischen Diversifikationsgrad und Performance. Während in entwickelten Volkswirtschaften ein negativer Zusammenhang zwischen Performance und Diversifikationsgrad besteht, ist dieser in weniger entwickelten positiv. Ähnlich mahnen auch Short et al. (2007) eine Einbeziehung geographischer Differenzierungen an. Auch Martin, Cullen, Johnson & Parboteeah (2007) finden landeskulturelle Faktoren als wesentliche erklärende Variable in ihrer Untersuchung zum Verhalten von Unternehmen zu Bestechungen. Takeuchi, Shay & Li (2008) kontrollieren in Ihrer internationalen Studie zum Verhalten von Expatriates für das Herkunftsland und finden teilweise hohe Signifikanzniveaus. Den Einfluss ordnungspolitischer Rahmenbedingungen, die für sich natürlich auch nicht unabhängig von der jeweiligen Landeskultur zu verstehen sind, untersucht etwa Capron & Guillén (2009). Die Autoren finden hierbei hochsignifikant unterschiedliche Möglichkeiten für Unternehmen, im Rahmen von Mergers Synergien zu realisieren.

Selbst innerhalb eines geographisch relativ begrenzten Bereiches kann es bereits substantielle Unterschiede geben. So beleuchtet Lounsbury (2007) das Verhalten von Investmentfonds in Boston und New York im Zeitraum von 1944 bis 1985 und findet substantiell unterschiedliche Managementlogiken. Starke Regionaleffekte auf das Überleben von (hier jungen) Unternehmen finden beispielsweise auch Fritsch, Brixy & Falck (2006). Auf die Relevanz dieses Faktors weist auch eine Studie von Hagedoorn & Cloudt (2003) hin. Die Autoren betonen die Notwendigkeit, zur Messung der Leistung angewandte Indikatoren auf ihre Eignung in jedem Sample zu hinterfragen. In einer ähnlich angelegten Studie finden Leiblein & Madsen (2009) dagegen keine statistisch signifikanten Einflüsse der Herkunft eines Unternehmens auf das (Innovations-) Ergebnis.

Zeitraum

Grundsätzlich müssen sich empirische Arbeiten auf einen bestimmten Betrachtungszeitraum beschränken, sei es ein Stichtag, wie im Falle einer Leistungsoperationalisierung mittels cumulative abnormal returns (CAR) oder bei einmaligen Fragebogenerhebungen notwendigerweise der Fall, seien es Monate, Jahre oder Jahrzehnte bei anderen empirischen Herangehensweisen. Die implizite Hoffnung ist, dass die Kausalitäten, die im betrachteten Zeitfenster oder zum betrachteten Zeitpunkt anzutreffen waren, sich auch in anderen Zeiträumen in gleicher Weise wiederfinden. Diese Hoffnung ist keineswegs trivial. Im erkenntnistheoretischen Kontext diskutiert insbesondere Gadamer (1990) die Historizität der Geistes- und Sozialwissenschaften (vergleiche Abschnitt 2.1). In der aktuellen praktischen Strategischen Managementforschung weist beispielsweise Bamberger (2008) auf die Bedeutung des Zeitpunktes als wesentlicher Kontextvariablen bei der Erarbeitung von Kausalaussagen hin. Grundsätzlich

sozialpsychologischen Perspektive (vergleiche etwa Scollon & Wong 2008). Eine detailliertere Betrachtung dieser Forschungsfelder liegt außerhalb des Fokus' dieser Arbeit, der interessierte Leser sei auf Publikationen wie die *Annals of Regional Science* oder das *Journal of Regional Science* verwiesen.

kann der betrachtete Zeitraum auf zweierlei Weg Einfluss nehmen. Zum einen im Rahmen kurzfristiger leistungsrelevanter Veränderungen im Umfeld, also insbesondere konjunkturelle Schwankungen oder Marktüberreaktionen („Blasen“), die in relativ kurz aufeinanderfolgenden Zyklen auftreten, zum Zweiten durch langfristige Veränderungen des Umfeldes, die etwa durch technischen Fortschritt oder gesellschaftliche Entwicklungen wie demographische Veränderungen oder Wertewandel hervorgerufen sein können.

Für die Relevanz konjunktureller Effekte liegt umfangreiche empirische Evidenz vor. Chakrabarti et al. (2007) identifizieren beispielsweise Zeit (wirtschaftlich stabil, wirtschaftlich turbulent) als wesentlichen erklärenden Faktor für den Zusammenhang zwischen Diversifikationsgrad und Performance, ähnlich auch Lee & Makhija (2009) oder Ngah-Kiing Lim, Das & Das (2009) in Untersuchungen zur Auswirkung wirtschaftlicher Schwächephasen. Sanders & Hambrick (2007) diskutieren die Begrenzung des Zeitraumes als ein Problem der Verallgemeinerbarkeit ihrer Ergebnisse zum Risikoverhalten von CEOs, ebenso Kroll, Walters & Le (2007) in ihrer Untersuchung von Top-Management Teams und ihrem Einfluss auf die Post-IPO Performance junger Unternehmen. Miller et al. (2007) kontrollieren in ihrer Studie zum Innovationsverhalten differenzierter Unternehmen für Zeiteffekte durch die Einbeziehung von Jahresdummies. Diese sind auf dem 5 %-Niveau statistisch signifikant, ähnlich auch bei Fitza, Matusik & Mosakowski (2009) in ihrer Studie zur Wertentwicklung von jungen Unternehmen. Die Autoren finden hier einen erheblichen Effekt aufgrund der dot.com-Blase von 1995-2000. Auch Kim & Finkelstein (2009) finden verschiedene signifikante Jahres-Dummies in ihrer Untersuchung zu Kontingenzzfaktoren bei Mergern.

Doch nicht nur konjunkturelle Zeiteffekte spielen eine Rolle, es scheint durchaus auch langfristig wirkende Tendenzen zu geben (vergleiche auch Robins, 2004). Ein aktuelles praktisches Beispiel für die Relevanz des betrachteten Zeitraums liefern Kronborg & Thomsen (2009). Die Autoren untersuchen den Zusammenhang zwischen dem Überleben von Unternehmen und der Frage, ob diese sich im Eigentum von Inländern, in diesem Fall Dänen, befinden oder ob es sich um Tochtergesellschaften ausländischer Unternehmen handelt. Interessant an dieser Studie ist insbesondere die Länge des betrachteten Zeitraums: Der verwendete Datensatz beginnt 1895 und erstreckt sich bis ins Jahr 2005. In der Untersuchung kann zunächst eine höhere Überlebenswahrscheinlichkeit der ausländischen Tochtergesellschaften ermittelt werden, diese nimmt jedoch über die Zeit ab und verschwindet in der letzten betrachteten Dekade vollständig. Die Beantwortung der Forschungsfrage hängt also zentral davon ab, welcher Ausschnitt aus dem weiten Zeithorizont der Studie Beachtung findet. Ein Datensatz, der etwa nur die 1970er oder 1980er Jahre berücksichtigt⁵⁰, hätte einen positiven Zusammenhang zwischen dem Überleben von Unternehmen und deren Status als Tochterunternehmen ausländischer Muttergesellschaften konstatiert, beginnt der Betrachtungshorizont in den 1990er-Jahren, gibt es keinen Zusammenhang. Es erscheint daher plausibel anzunehmen, dass eine entsprechende zeitliche Eingrenzung der Grundgesamtheit in die Interpretation etwaiger Forschungsergebnisse zentral und relativierend einbezogen werden müsste. Oder, populär ausgedrückt: Es gibt keinen Grund anzunehmen, dass das, was 1980 oder 1990 richtig war, auch 2010 oder 2020 noch richtig sein muss. Diese Relativierung ist eine Aufgabe, die im Rahmen der Weiterverwendung von Wissen zu lösen ist, das in vergangenen Perioden

⁵⁰ Dieser Zeithorizont wird beispielsweise in den aktuellen Studien von Henkel (2009), 1970-1979, Koka & Prescott (2008), 1980-1994 oder Sinha & Noble (2008), 1981-1986, verwendet, um nur einige wenige Beispiele zu nennen. Vergleiche zum Thema der Alterung von Datensätzen auch Robins (2004).

gewonnen wurde. Hier gilt es, eine begründete Annahme darüber zu treffen, inwieweit sich relevante Rahmenbedingungen geändert haben. In obigem Fall führen die Autoren das Verschwinden des Überlebensbonus auf den verschärften Wettbewerb zwischen den ausländischen Tochtergesellschaften zurück, der sich mit dem sich seit einigen Jahrzehnten stark verschärfenden internationalen Wettbewerb in Zusammenhang bringen lässt (Kronborg & Thomsen, 2009). Auch Shimizu (2007) findet in seiner Untersuchung zu Desinvestment-Entscheidungen nach erfolgten Mergern, dass sich die Wahrscheinlichkeit zum Desinvestment übernommener Unternehmen bis maximal 10 Jahre nach erfolgter Akquise über zwei Dekaden substantiell geändert hat.

Notierung und Größe

Die Verwendung marktbasierter Leistungsindikatoren verhindert die Betrachtung nicht börsennotierter Unternehmen, der wesentlichen Mehrzahl aller Unternehmen einer Volkswirtschaft (siehe oben). Dieses Problem wird beispielsweise in Tong, Reuer & Peng (2008) gesehen. Macher & Boerner (2006) kontrollieren in ihrer Studie zum Zusammenhang zwischen dem Lernverhalten von Unternehmen, Skalen- und Fokussierungseffekten und der Performance in der Produktentwicklung spezifisch für die Notierung der Unternehmen. Die Autoren finden einen signifikanten Unterschied zwischen gelisteten und nicht gelisteten Unternehmen. Dagegen scheint für die Marktzutrittsstrategien in der Biotech-Branche die Notierung der Unternehmen keine Rolle zu spielen (Aggarwal & Hsu, 2009). Einen signifikanten Einfluss der Notierung findet auch Ang (2008). In dieser Arbeit wird der Zusammenhang zwischen Wettbewerbsintensität und Kooperationsneigung auf der einen Seite und dem Wachstum von Unternehmen auf der anderen Seite über verschiedene technologische Umwelten hinweg untersucht.

Die Größe eines Unternehmens hat einen wesentlichen Einfluss auf Strukturen und Hierarchien, Risikowahrnehmung und vieles andere mehr (vergleiche beispielsweise Williamson, 1999; Iyer & Miller, 2008). Daher verwundert es nicht, dass eine Verallgemeinerbarkeit über unterschiedliche Größenklassen hinweg zumindest als problematisch angesehen wird. Iyer & Miller (2008) beispielsweise finden einen signifikanten Einfluss der Unternehmensgröße auf das Akquisitionsverhalten von Unternehmen. Die Einbeziehung der Unternehmensgröße als Kontrollvariable ist in vielen Untersuchungen Standard (beispielsweise Mesquita & Lazzarini, 2008; Canella Jr. et al., 2008; Weigelt, 2009; Levitas & McFayden, 2009), die statistische Signifikanz schwankt je nach betrachteter Fragestellung, ist aber häufig gegeben.

Datenbanken

Ein besonderer Fall von Selektionsfilter sind Datenbanken. Der Zugriff auf bereits aufbereitete Sekundärdaten stellt, neben der eigenen Erhebung mittels Fragebogen und der Auswertung von Unternehmens- oder Pressepublikationen, den dritten wesentlichen Zugangsweg zu Unternehmensdaten für Forscher in der Strategischen Managementforschung dar. Durch den Rückgriff auf Sekundärdaten einer bestimmten Datenbank übernimmt der Forscher allerdings die Vorselektion, die dieser Datenbank zu Grunde liegt, sowie alle Fehler, die in den Daten

möglicherweise enthalten sind.⁵¹ Dies kann die Verallgemeinerbarkeit der gewonnenen Ergebnisse beschränken, abhängig von der Fallselektion und der Datenqualität der genutzten Datenbank. Die Relevanz dieses Problems untersucht etwa Schilling (2009). Die Autorin geht der Frage nach, ob das Problem der Populationsselektion sich auf die Ergebnisse von Studien auswirkt, die bestimmte Datenbanken als Grundlage nutzen, beziehungsweise ob sich andere Ergebnisse einstellen würden, würde man eine andere Datenbank zugrunde legen. Schilling vergleicht vier Datenbanken, die zur Auswahl der Untersuchungsobjekte häufig herangezogen werden, und rechnet drei publizierte Studien, die sich auf bestimmte Datenbanken zur Auswahl der Forschungsobjekte stützen, mit anderen, inhaltlich dasselbe erklärenden Datenbanken neu. Die Autorin findet lediglich in einem von drei untersuchten Fällen signifikante Abweichungen, wenn eine andere Datenbank und damit eine andere Teilpopulation Untersuchungen zu Grunde gelegt wurde, so dass die Autorin einen eher geringen Einfluss dieses Populationsselektionsinstruments vermutet. Dem steht jedoch auch ihre Beobachtung gegenüber, dass die Datenbanken selbst auf ähnlichen Selektionsmustern basieren (in diesem Fall: größere Unternehmen aus dem OECD-Raum mit hoher Frequenz an Allianz-Vereinbarungen), was die Aussage wiederum relativiert. Entsprechend warnt Schilling vor einer Verzerrung, wenn etwa kleinere oder jüngere Unternehmen im Mittelpunkt des Interesses stehen. Auch ein Einfluss der Bekanntheit eines Unternehmens auf seine Aufnahme in eine der untersuchten Datenbanken wird vermutet. Allgemein lässt sich die Selektionswirkung von Datenbanken minimieren, wenn einer Studie mehrere, im Bereich der Selektion der aufgenommenen Objekte möglichst heterogene Datensammlungen zu Grunde gelegt werden, wie beispielsweise in Pacheco-de-Almeida, Henderson & Cool (2008) oder Stam & Elfring (2008), dies erfordert allerdings natürlich deren Existenz und eine gewisse grundlegende Homogenität in den Items, die die jeweiligen Datenbanken zu liefern in der Lage sind.⁵²

3.2.4 Diskussion

Es ist also festzuhalten, dass bei allen der häufig verwendeten Eingrenzungen der Grundgesamtheit eine Verallgemeinerbarkeit über die Grenzen der betrachteten Teilpopulation hinweg regelmäßig nicht zu erwarten ist. Die Heterogenität von Forschungsergebnissen auf unterschiedliche Teilpopulationen zurückzuführen ist damit also konzeptionell folgerichtig. Was folgt hieraus für die Methodik der Strategischen Managementforschung? Zuvorderst Skepsis gegenüber der grundsätzlichen Möglichkeit zu objektiven, das heißt verallgemeinerbaren, Ergebnissen zu kommen. Die Selektion einer Teilpopulation kann ein wesentlicher Schritt sein, verschiedene „Proverbs“ in der Sprachregelung von Simon (1946), Komplexitätserhöhende Variablen, in die Untersuchung einzubeziehen oder auszuschließen. Umso mehr ent-

⁵¹ Diese Fehler sind laut Starbuck (2004) teilweise erheblich. Insbesondere für die Compustat-Datenbank, die zahlreichen Studien zu Grunde liegt, konstatiert er einen sehr hohen Anteil falsch eingegebener Einträge, was die Ergebnisse von Studien, die auf diesen Datensätzen aufbauen, massiv verzerren kann. Entsprechend kritisch äußert sich Starbuck zur wissenschaftlichen Verwertung von Sekundärdaten aus großen Datenbanken.

⁵² Vergessen werden darf an dieser Stelle auch nicht, dass die Alternative zur hier geschilderten Arbeit mit Sekundärdaten aus Datenbanken, die Durchführung eigener Erhebungen, also die Verwendung von Primärdaten, ebenfalls mit erheblichen Selektionsproblemen behaftet ist, wie etwa Slater & Athuahene-Gima (2004) ausführlich diskutieren. Auch hier ist die Repräsentativität der gezogenen Stichprobe für bestimmte Grundgesamtheiten ein zentrales Problem, die Auswahl der Adressaten ist schwierig und ein Fallstrick für Biases. Non-Response-Biases können weitere Verzerrungen in die Studie einbringen.

steht hierdurch aber die Notwendigkeit, die aus einem bestimmten Studienaufbau gewonnenen Erkenntnisse kritisch vor dem Hintergrund der jeweils untersuchten Teilpopulation zu diskutieren und die Einflussfaktoren, die aus dieser spezifischen Population herrühren und die Ergebnisse beeinflussen (können) als solche zur Kenntnis zu nehmen beziehungsweise zu bringen, wie etwa auch von Short und Kollegen (2002) dringend nahegelegt.

Im Rahmen der Meta-Analyse in Kapitel 5 wird zu eruieren sein, inwieweit das Problem der heterogenen Teilpopulationen sich systematisch auf die Ergebnisse wissenschaftlicher Arbeiten auswirkt und die Heterogenität in den Ergebnissen der Strategischen Managementforschung erklären kann. Wie oben gezeigt wurde, legt sowohl die wissenschaftstheoretische Perspektive als auch ein Blick in die Literatur zu den relevantesten Populationsselektionskriterien diese Annahme nahe. Die Begrenzung eines Studienaufbaus auf eine Teilpopulation der Gesamtpopulation aller Unternehmen führt zu einer Reduktion der Erklärungsmacht von in dieser Teilpopulation durchgeführten Hypothesentests. Verschiedene Autoren bestätigten in wissenschaftlichen Arbeiten die Bedeutung von Branche, regionaler Herkunft, Notierung und Größe von Unternehmen für Arbeiten in der Strategischen Managementforschung. Zu beachten ist auf der anderen Seite jedoch, dass die Fokussierung auf Teilpopulationen durch die damit einhergehende Ausschaltung möglicher Moderatorvariablen als notwendiger Schritt auf dem Weg zu statistisch belastbaren Ergebnissen zu sehen ist. In der in Kapitel 4 folgenden Literaturreview wird zu klären sein, welche Populationsselektionskriterien wie häufig Verwendung finden. Die Metaanalyse in Kapitel 5 schließlich wird überprüfen, inwieweit die Untersuchung unterschiedlicher Teilpopulationen Heterogenität in den Forschungsergebnissen im Fachgebiet der Strategischen Managementforschung erklärt.

3.3 Das Appropriationsproblem

Ein zentrales Problem der Leistungsmessung besteht darin, dass mit der Wahl einer bestimmten Performanceoperationalisierung implizit auch eine Wertung über die erwünschte Zielfunktion der untersuchten Organisation einhergeht. Diese wiederum ist eng mit der Theorie der Organisation verwoben. Das Bild, das der Wissenschaftler (oder Manager, Analyst etc.) von der Theorie einer Organisation explizit oder implizit vor Augen hat, beeinflusst sein Denken und Urteilen in vielerlei Dimensionen (vergleiche Zingales, 2000; Seisreiner, 2006; Asher, Mahoney & Mahoney, 2005; Goshal, 2005). Entsprechend ist es an dieser Stelle notwendig, die Beziehung zwischen Organisationstheorie und Leistungsmessung zu beleuchten. Ich wähle hierzu den Ansatz der Neuen Institutionenökonomik (vergleiche auch Wolf, 2008), und dessen Kritik, beziehungsweise Weiterentwicklung, die für die folgenden Überlegungen zur Zielfunktionen von Organisationen eine hohe Erklärungsrelevanz aufweisen. Aufbauend auf der Agency-Theorie, die die Unternehmung als Nexus expliziter Verträge begreift, und der Kritik an diesem Ansatz diskutiere ich die jeweiligen Implikationen für die Zielfunktion und in der Folge Performance-Messung und kontrastiere sie mit Ansätzen der Stewardship- und Stakeholder-Theorie, die als Kritik am Agency-Ansatz verstanden werden können. Aus der Agency-Theorie lässt sich eine konzeptionelle Rechtfertigung für eine auf den Shareholder Value aufbauende Zielfunktion von Unternehmen ableiten. Wäre diese Annahme zu halten, wäre das Problem der Performanceoperationalisierungen wesentlich vereinfacht: Es ginge darum, den Shareholder Value möglichst exakt zu bestimmen. Ich gehe von der Arbeitshypothese aus, dass der konzeptionellen Grundlagedes Shareholder Value-Ansatzes eine zu stark vereinfach-

te Sichtweise der Organisation zu Grunde liegt, die für Erkenntnisfortschritte im Fachgebiet der Strategischen Managementforschung dysfunktional wirken kann. Wäre dem so, wäre die „natürliche“ Fokussierung auf den Shareholder Value in der Leistungsoperationalisierung nicht mehr gegeben, die Wahl der Operationalisierung wäre, wenn sie eine implizite Annahme über die Zielfunktion eines Unternehmens beinhaltet, begründungsbedürftig. Im Zuge dieser Begründung gilt es, ein Verständnis der betrieblichen Wertschöpfung zu schaffen und die Verteilung dieser Wertschöpfung in der Organisation zu diskutieren. Dies führt zu weiteren Betrachtungen auf Basis von Stakeholder- und Appropriationstheorie und schlussendlich zur Ableitung einer Hypothese über ein mögliches systematisches Methodenproblem im Zusammenhang mit der Leistungsoperationalisierung in der Strategischen Managementforschung.

3.3.1 Organisationstheoretische Grundlagen

Eine Organisation besteht, auch wenn sie als eigenständige Rechtspersönlichkeit existiert und auch in der Forschung häufig aggregiert betrachtet wird, immer aus Individuen, die sich aus bestimmten Gründen dazu entschieden haben, gemeinsam diese Organisation zu bilden, beziehungsweise eine Beziehung zu einer durch andere Individuen gebildeten Organisation einzugehen. Die Betrachtung einer Organisation als homogene Einheit stellt eine Vereinfachung dar, die, wenn auch häufig notwendig und gerechtfertigt, immer als solche verstanden werden muss (Felin & Foss, 2005). Ein Unternehmen entsteht, wenn für die Teilnehmer, die am Unternehmen partizipieren, die Kosten, die durch die Partizipation entstehen, also beispielsweise die Kosten der Lieferung für Lieferanten, das „Arbeitsleid“ für Mitarbeiter und so weiter, geringer ausfallen als der durch den impliziten oder expliziten Vertrag mit dem Unternehmen jeweils generierte Nutzen, also die Vergütung für den Lieferanten, die verschiedenen Incentives für den Mitarbeiter und so weiter (Simon, 1952).⁵³ Hierbei stellt sich in der statischen Betrachtung ein paretooptimaler Zustand in einer einmal gebildeten Organisation ein, in dem alle Partizipanten in der Organisation einen individuellen Nutzenwert von größer 0 erzielen, das heißt sie erhalten in ihrer jeweils individuellen Bewertung mehr Anreize die Beziehung mit der Organisation aufrechtzuerhalten als diese abubrechen, der Nutzen übertrifft die Kosten. Wäre dem nicht so, hätten sie ihre Vertragsbeziehungen mit der Organisation beendet, beziehungsweise wären eine solche nie eingegangen. Die Kosten umfassen hier neben den tatsächlichen Kosten des Inputs auch die Opportunitätskosten eines alternativen Kontrakts (Simon, 1952), also beispielsweise Chancen und Risiken im Falle eines Arbeitgeberwechsels für Mitarbeiter, die Möglichkeiten zum Verkauf der Vorprodukte an andere Abnehmer beziehungsweise den Verlust eines Kunden für Lieferanten und so weiter.

⁵³ In seinem grundlegenden Artikel hierzu bezeichnet Simon, 1952, 41f den Rückfluss der Stakeholder als „inducement“, ihren Input als „contribution“. Die zentrale implizite Annahme von Simons Analyse ist, dass alle Akteure in der Wahl ihrer Vertragspartner frei sind, es also keine Verpflichtung gibt, einen Vertrag mit einer bestimmten Institution oder Person abzuschließen oder einer bestimmten Institution anzugehören. Diese Annahme ist nicht immer haltbar, Gegenbeispiele wären etwa der Fall eines Zivildienstleistenden in einem Pflegebetrieb, oder Unternehmen, die einem hoheitsrechtlich verordneten Kontrahierungszwang unterliegen (etwa gesetzliche Krankenversicherungen). In weiten Teilen des wirtschaftlichen Lebens kann jedoch von der Existenz einer solchen Vertragsfreiheit grundsätzlich ausgegangen werden.

Wichtig ist hierbei die Annahme begrenzt opportunistischen Verhaltens⁵⁴. Die Vertragsparteien gehen Beziehungen und Handlungen zwar nicht aus idealistischen sondern aus eigennutzgeleiteten Gründen ein und versuchen im Rahmen dieser Verträge, in einer dynamischen Betrachtung, ihre jeweilige Individualrente zu maximieren. Gleichzeitig bringen sich die Vertragsparteien jedoch gegenseitig so viel Vertrauen entgegen, dass auch unvollständig spezifizierte Verträge möglich bleiben und auf eine Optimierung des Individualnutzens um den Preis einer erheblichen Schädigung der anderen Vertragsparteien bewusst verzichtet wird. Kooperation stellt also für die Vertragsparteien einen für sich nutzenstiftenden Wert dar (Davis, Schoorman & Donaldson, 1997). Die Parteien agieren demnach grundsätzlich wohlwollend gegenüber der Organisation, mit der sie durch implizite oder explizite Verträge verbunden sind, ohne radikal ihren Individualnutzen zu maximieren. Insbesondere das Management agiert als „steward“, als wohlmeinender Verwalter im besten Interesse etwa der Kapitalgeber, ohne dass hierfür exzessive Kontrollmechanismen notwendig wären. Der Steward bewertet im Gegensatz zum Agenten der Principal-Agent-Theorie (siehe unten) Kooperation für sich als nutzenstiftender als Defektion (Donaldson & Davis, 1991; Davis et al., 1997). Durch das auf diese Weise geschaffene Vertrauen werden für die Vertragsparteien auch spezifische Investitionen in das Unternehmen rational, sunk costs werden akzeptabel (Asher et al., 2005). Die Verfolgung individueller Interessen bleibt hiervon grundsätzlich unberührt, der Stakeholder versucht also im Rahmen der Regelungen seines Vertrages mit der Organisation weiter, für sich Incentives zu generieren: der Lieferant strebt nach besten Preisen, der Arbeitnehmer nach hohem Gehalt und sozialer Absicherung und so weiter (vergleiche hierzu auch Williamson, 1999; Goshal, 2005)⁵⁵, allerdings nicht um den Preis massiver Schädigung der Vertragsparteien und einer Zerstörung des Vertrauensverhältnisses.

Mit diesem relativ positiven Menschenbild grenzt sich die Stewardship-Theorie von der in der Strategischen Managementforschung weiter verbreiteten Agency-Theorie (Sundaram & Inkpen, 2004b; Jensen, Meckling, Field & Park, 1976; Plender, 1998) ab, die von einem individualnutzenmaximierenden Homo Öconomicus ausgeht. In dieser Theorie übertragen die Eigentümer die ihnen zustehende Macht über das Unternehmen als Prinzipale an Agenten, das Management, das nun wiederum bestrebt ist, diese Macht zu seinem eigenen Vorteil und damit potentiell zum Schaden der Prinzipale einzusetzen, um wiederum seinen eigenen Individualnutzen zu maximieren. Um die Interessen der Prinzipale zu schützen müssen in diesem Verständnis Mechanismen entwickelt werden, die dafür sorgen, dass die Agenten gezwungen werden, so weit wie möglich im Interesse der Prinzipale zu agieren. Ein System vollständig explizierter Verträge ist hierfür notwendig, das Rechte und Pflichten der Vertragspartner exakt festlegt und somit die Defektion der Agenten unterbindet. Wesentliche Instrumente hierzu sind etwa Entlohnungsformen, die den Individualnutzen der Agenten mit denen der Prinzipale in Einklang bringen sollen (Davis et al., 1997; Castanias & Helfat, 2001). Die Agency-Theorie führt zwangsläufig zu einer Ausrichtung der Unternehmensziele auf die Interessen der Anteilseigner hin, denen als Prinzipale axiomatisch die Herrschaft über die Verfügungsrechte zugeschrieben wird und die vor der Willkür des Managements zu schützen sind (Asher

⁵⁴ „bounded self interest“, vergleiche Bosse, Phillips & Harrison, 2009, 447.

⁵⁵ Analog zum Konzept des Shareholder Value spricht man in diesem Zusammenhang dann gelegentlich von „employee value“, „customer value“ oder „management value“ (Payne, Holt & Frow, 2000; Seisreiner, 2006).

et al., 2005; Seisreiner, 2006).⁵⁶ Durch diesen Fokus auf den expliziten Vertrag zwischen Anteilseignern und Management definiert die Agency-Theorie die Organisation als Nexus expliziter Verträge zwischen Prinzipalen und Agenten (Zingales, 2000). Der Agentenstatus pflanzt sich hierbei in der Hierarchie nach unten fort, untergeordnete Führungskräfte agieren als Agenten des Top-Managements und so weiter.

Im Gedankengebäude der Agency-Theorie delegieren die Anteilseigner, die Prinzipale, also Macht an ihre Agenten, das Management, das den Zielen der Prinzipale dienend mit dieser umzugehen hat. Wenn die logische Schlussfolgerung ist, dass die Prinzipale vor der Willkür und dem Machtmissbrauch der Agenten geschützt werden müssen ist die Leistung einer Organisation dieser Logik folgend daran zu messen, wie gut sie den Zwecken der Prinzipale gedient hat. Agenten sind dann positiv zu bewerten, wenn sie die Macht, die ihnen überlassen wurde, nicht defätistisch genutzt sondern dem Zweck des Prinzipals dienlich eingesetzt haben. Es ergibt sich somit ein klares Primat der Anteilseigner, die Zielfunktion für das Unternehmen ist klar: Es hat den Zwecken der Prinzipale zu dienen. Zahlreichen Arbeiten in der Strategischen Managementforschung liegt implizit dieses Shareholder-Paradigma zu Grunde⁵⁷, wonach der Zweck einer jeden Unternehmung in der Generierung von Werten für die Eigenkapitalgeber liegt. Alle anderen Ziele haben sich diesem Oberziel unterzuordnen oder sind dahingehend zu prüfen, inwieweit sie effektiv und effizient zur Erreichung des Oberziels wirken. Atkinson, Waterhouse & Wells (1997) bezeichnen das Shareholder Value-Ziel diesem Gedanken folgend entsprechend als „primary“, alle weiteren Ziele als „secondary objectives“ (S. 31).

In einer neueren Erweiterung, dem Stewardship-Ansatz, wird Managern, wie erwähnt, eine wohlwollendere Haltung zum Unternehmen zugestanden (Jensen et al., 1976; Asher et al., 2005; Wasserman, 2006), grundlegend bleibt jedoch, dass der Leistungserstellungsprozess und die Wertschöpfung eines Unternehmens im Wesentlichen nur zwischen den beiden Gruppen Manager und Eigentümer diskutiert wird und andere Anspruchsgruppen allenfalls als Produktionsfaktoren oder Größen der Produktionsfunktion in Erscheinung treten. Über die Vertretbarkeit dieser Bevorzugung der Gruppe der Anteilseigner gegenüber anderen Stakeholdern tobt ein erbitterter Streit in der Managementforschung, in den einzusteigen an dieser Stelle aufgrund seines weniger wissenschaftlich als vielmehr ethisch wertend aufgeladenen Charakters nicht zielführend ist, der interessierte Leser sei beispielsweise auf Sundaram & Inkpen (2004a), Freeman, Wicks & Parmar (2004), Sundaram & Inkpen (2004b), Heugens & Lander (2009) und Donaldson & Preston (1995) verwiesen. Fraglich ist, ob es sich bei der zu beobachtenden Auseinandersetzung tatsächlich um einen wissenschaftlichen Wettstreit zwischen empirisch überprüfbaren Theorien im Popperschen Sinne (vergleiche Abschnitt 2.1) handelt, oder nicht vielmehr um einen wissenschaftlich gefärbten Austausch präskriptiver Gedankengebäude, weniger ein „was ist“ als vielmehr ein „was sollte sein“. Die weitgehende Irrelevanz für die hier zu analysierende Fragestellung ist auch dadurch begründet, dass es im Wesentlichen um normative Aussagen für die Unternehmensführung und der

⁵⁶ Eine ausführliche Darstellung der Probleme der Principal-Agent Beziehung und der Gefahren für den Prinzipal findet sich beispielsweise bei Picot, Dietl & Franck (2008).

⁵⁷ Vergleiche Abschnitt 4.2.2: Etwa die Hälfte aller Studien arbeitet mit Performanceoperationalisierungen, die die Leistung des Unternehmens aus Sicht der Anteilseigner messen: marktbasierende und accountingbasierte Kenngrößen.

Abwägung von Interessen innerhalb einer Organisation geht, weniger um ein Verständnis der Leistung dieser Organisation.⁵⁸

Die Agency-Theorie wird häufig aus Gründen kritisiert, die auch für den Argumentationsgang dieser Arbeit von Relevanz sind. Der besondere Schutz, den die Eigentümer deshalb genießen, weil sie aufgrund ihres Prinzipal-Status und der damit unter anderem einhergehenden schlechten Informationsversorgung am leichtesten zu schädigen sind, erscheint fragwürdig, da durch Entscheidungen des Managements auch andere Personen, etwa Arbeitnehmer, unmittelbar negativ betroffen sein können, häufig in einer deutlich schwereren persönlichen Konsequenz als dies bei Anteilseignern der Fall wäre (Zingales, 2000; Asher et al., 2005; Seisreiner, 2006). Mit dem Fokus auf lediglich zwei Vertragsparteien vereinfacht die Agency-Theorie den Nexus von Verträgen, der eine Organisation bildet, in einer Weise, die die Analysetauglichkeit der Theorie für viele Situationen stark begrenzt (Zingales, 2000). Die Unmöglichkeit vollständig explizierter Verträge wird nicht beachtet und sunk costs, die ein Organisationsmitglied auf sich nimmt, sind schwer erklärbar (Lorenz, 1999; Tirole, 1999). Schließlich wird das grundlegende Axiomensystem der Agency-Theorie als zu simplifizierend kritisiert (Davis et al., 1997; Zingales, 2000).

In der Agency-Theorie sind Verträge, die nicht vollständig expliziert sind, grundsätzlich schädlich, bilden sie doch Einfallstore für opportunistische Handlungen zum Schaden der Organisation und damit der Anteilseigner. Dies bedeutet auch, dass alle Handlungen, die Akteure zum Wohle und damit als Beitrag zur Leistung einer Organisation erbringen, vorab vertraglich spezifiziert sind, beziehungsweise spezifiziert werden müssen. Kein Akteur wird mehr einbringen, als er vertraglich einzubringen verpflichtet ist, der Vertrag kontrolliert, dass er auch nicht weniger einbringt. Die Leistung eines Unternehmens lässt sich demzufolge als Summe der vertraglich vereinbarten Leistungen exakt definieren (Zingales, 2000). Dies ist offensichtlich für viele Fragen der Strategischen Managementforschung eine unzulässige Vereinfachung, gäbe es hier doch wenig Spielraum für nachhaltige, nicht erklärte Leistungsunterschiede zwischen Organisationen, die auf endogenen, also organisationsinternen Faktoren beruhen. Unterschiede in der Leistung von Organisationen wären demnach nur durch exogene Faktoren, etwa Umweltveränderungen, exogene Schocks, zu erklären.

⁵⁸ Vergleiche hierzu beispielsweise den Diskurs dazu, ob es sich für Unternehmen „lohnt“, sich sozial zu engagieren: Karnani, 2007; Godfrey, Merrill & Hansen, 2009; Barnett, 2007. Aus Sicht der Agency-Theorie ist diese Frage von Interesse, geht man aber auf eine Theorie impliziter Verträge, von Stewardship und Stakeholdern (siehe unten) über, ist diese Fragestellung redundant, da sie, wie auch Kochan & Rubinstein (2000) anmerken, versucht, eine Abwägungsentscheidung des Managements mit wissenschaftlichen Methoden zu bewerten, beziehungsweise zu rationalisieren (auch: Asher, Mahoney & Mahoney, 2005). Normativer Empfehlungen muss sich die Wissenschaft aber enthalten, so lange keine Theorien existieren, die Kausalbeziehungen aufzeigen können, die eine besser/schlechter Wertung ermöglichen. Bei der Frage nach Stakeholder/Shareholder-Orientierung kann die Strategische Managementforschung bisher allenfalls die Funktionalität bestimmter Faktoren und Gegebenheiten für bestimmte Zielfunktionen beschreiben; die Definition dieser Zielfunktionen jedoch ist nicht Gegenstand wissenschaftlichen Erkenntnisstrebens sondern von Werturteilen oder allenfalls noch wirtschaftsphilosophischer oder wohlfahrtsökonomischer Gerechtigkeitstheorien (vergleiche beispielsweise Blankart, 2008, Kapitel 5; Goshal, 2005). Dem Ideal einer weitgehend werturteilsfreien Wissenschaft, die aufdecken soll, was ist, kann man sich auf diese Art und Weise nicht nähern (vergleiche Kapitel 2.1). Oder, wie Goshal, 2005, 86, es formuliert: „Wishes and hopes are not theory. Sermons and preaching are not theory either“.

Die Erweiterung dieser Theorie hin auf die Berücksichtigung expliziter und impliziter Verträge lässt dagegen, durch den impliziten Charakter der Vertragsgestaltung, Raum für gewachsenes Vertrauen und die Unmöglichkeit, alle Leistungen vorab zu spezifizieren, und ermöglicht es so, Abweichungen von der im Vertrag spezifizierten Leistung der Vertragspartner anzunehmen.⁵⁹ Nur auf diese Weise kann der Wert, das heißt die Leistung einer Organisation, tatsächlich unklar sein. In der Agency-Theorie ist die Leistung einer Organisation immer exakt die Summe ihrer Teile, im Gedankengebäude der impliziten Verträge dagegen hängen Wert und Leistung einer Organisation von individuellen, schlecht spezifizierten und spezifizierbaren Inputs der Vertragsparteien ab. Diese können zusätzliches Kapital verschiedenster Art einbringen oder auch defätistisch handeln, womit die Leistung eines Unternehmens nach oben oder unten von der in Verträgen explizierten Leistung abweicht (Zingales, 2000).⁶⁰ Nur in diesem Gedankengebäude sind etwa die neueren Entwicklungen des Resource-based view und auch Abspaltungen wie der Competence-based view oder der Dynamic Capabilities Approach denkbar, die von einem wesentlichen Übersetzungsschritt zwischen Ressourcen und Wettbewerbsvorteil ausgehen (Peteraf & Barney, 2003; Ray et al., 2004; Schreyögg & Kliesch-Eberl, 2007). In einer Welt ausschließlich explizierter Verträge würden Ressourcen tatsächlich automatisch zu Wettbewerbsvorteil und Erfolg führen, der Tautologievorwurf, der dieser Theorie häufiger vorgehalten wurde (Powell, 2001). Erst die Ambiguität, die mit der Annahme impliziter Verträge einhergeht, ermöglicht konzeptionell eine Annäherung an das von Rumelt et al. (1994) so genannte „problem of persistent profits“. Die simplifizierenden Annahmen der Agency-Theorie scheinen daher als Grundlage der Analyse vieler Fragestellungen der Strategischen Managementforschung ungeeignet, insbesondere für solche, die auf Organisationen oder Prozesse abzielen, die unklar abgegrenzt oder schwer zu spezifizieren sind, etwa im Fall von Mitarbeiterkompetenzen, prozeduralem Wissen oder organisationaler Anpassungsfähigkeit (Teece, 2000; Haas & Hansen, 2007).

In der Diskussion um das angemessene Menschenbild für die Managementpraxis kann es dagegen, wie Davis et al. (1997) ausführlich diskutieren, keine entweder-oder-Abwägung zwischen Agency- und Stewardship-Theorie beziehungsweise Menschenbild geben.⁶¹ Ob Manager als Agenten oder Stewards handeln hängt vielmehr von verschiedenen psychologischen, situativen und unternehmenskulturellen Aspekten ab. Diese Erkenntnis unbenommen bleibt für den Argumentationsstrang dieser Arbeit jedoch mit Zingales (2000) festzuhalten, dass ein aus der Agency-Theorie abgeleitetes Unternehmensverständnis zwangsläufig zu ei-

⁵⁹ Vergleiche zur Ende der 1990er-Jahre unter anderem aus der Unzufriedenheit mit den Annahmen der Principal-Agent-Theorie (Tirole, 1999) heraus entstandenen Theorie unvollständiger Verträge, die für folgenden Betrachtungen die Basis liefert, und die Rolle von Vertrauen in der Vertragsgestaltung die zentralen Beiträge von Hart & Moore (1999), Lyons & Mehta (1997), Lorenz (1999) und Baker, Gibbons & Murphy (2002).

⁶⁰ Relevant ist hier insbesondere die Einbringung von Wissenskapital, das naturgemäß schwer exakt fassbar und damit nicht in expliziten Verträgen abbildbar ist, vergleiche Pawlowsky, Reinhardt, Bornemann & Schneider (2003).

⁶¹ Die empirische Evidenz für beide Ansätze ist, diese Ansicht bestätigend, auch relativ gemischt, vergleiche neben Davis, Schoorman & Donaldson (1997) auch Demsetz & Lehn (1985), Wasserman (2006), Makri, Lane & Gomez-Mejia (2006) und Jones, Felps & Bigley (2007). Vergleiche aber auch Goshal (2005), der von einer Art selbsterfüllenden Prophezeiung des negativen Menschenbildes durch entsprechendes Priming von Managern in der Ausbildung ausgeht. Goshal diskutiert auch erhellend die Diffusion des negativen Menschenbildes aus der Philosophie des Radikalindividualismus durch die Soziologie hin in die Managementwissenschaften und die daraus resultierenden negativen Menschenbilder, Kontrollerfordernisse und Rechtfertigungen a- und unmoralischen Verhaltens.

ner Leistungsdefinition führt, die durch die Fokussierung auf lediglich zwei Stakeholder-Gruppen stark realitätsverkürzend für Fragen der Strategischen Managementforschung wirken kann.

Eine konzeptionelle Abkehr von der Agency-Theorie hat allerdings eine fundamentale Implikation für die Operationalisierung von Leistung: Sie beraubt die Shareholder-Value-Sichtweise ihrer zentralen theoretischen Grundlage. In einer Welt impliziter und expliziter Kontrakte, multipler Verbindungen, von Vertrauen und Abhängigkeiten, in denen Leistungen und Pflichten nicht vollständig explizierbar und kontrollierbar sind, entfällt die allgemeingültige Rechtfertigung, einen „letztgültigen“ Prinzipal, den Anteilseigner, in der Zielfunktion eines Unternehmens⁶² herauszuheben. Eine solche Heraushebung ist natürlich im Rahmen unternehmerischer Entscheidungen oder auch im wissenschaftlichen Kontext weiterhin möglich, erfordert nun aber eine explizite Begründung oder eine organisationspezifische Abwägungsentscheidung. Entsprechend der oben erwähnten Kritik von Zingales (2000) können in einer Welt expliziter und impliziter Verträge Schädigungen auch auf völlig anderen Ebenen als nur zwischen Eigentümern und Beauftragten auftreten, entsprechend sind alle Vertragsparteien potentiell Opfer oder Täter schädigenden Verhaltens. Die Anteilseigner werden so zu einer unter mehreren Gruppen relativiert. In Konsequenz ist die exklusive Orientierung der Zielfunktion von Organisationen am Shareholder Value und die Abwendung von der Stakeholder-Orientierung, die seit Jahrzehnten in der Praxis zu beobachten ist (Plender, 1998), konzeptionell nicht mehr zwingend. Die Strategische Managementforschung steht in der Folge in einer Begründungspflicht im Rahmen der Operationalisierung von Leistung. Eine natürliche Orientierung an den Anteilseignern, wie sie etwa bei Performanceoperationalisierungen aus dem accounting- oder marktorientierten Bereich zu konstatieren ist, wird erklärungsbedürftig.⁶³

Damit ist für jede Performanceoperationalisierung zu hinterfragen, ob sie tatsächlich geeignet ist, die Leistung einer Organisation zu messen. Wie ausführlich in den Abschnitten 2.1 diskutiert, muss eine sinnvolle Leistungsoperationalisierung logisch aus dem wissenschaftlichen Kontext heraus erfolgen und insbesondere der Forderung nach konvergenter und diskriminanter Validität genügen. Dies bedeutet, dass Operationalisierungen, die Unterschiedliches messen sollen, tatsächlich Unterschiedliches messen (problematisch etwa im Rahmen der Unterscheidung zwischen Wettbewerbsvorteil und Performance) und Konstrukte, die dasselbe messen sollen, tatsächlich dasselbe messen. Letzteres ist im Bereich der Performanceoperationalisierung der kritische Faktor. Leistungsoperationalisierungen sollen die Leistung einer Organisation möglichst robust messen. Fraglich ist, ob die Forderung nach

⁶² Vergleiche zum Thema institutioneller Legitimation von Strukturen und Zielen in Organisationen auch DiMaggios und Powells viel zitierten Artikel von 1983 (DiMaggio & Powell, 1983), in dem sie das Phänomen des institutionellen Isomorphismus diskutieren. Organisationen tendieren demnach dazu, sich anderen Organisationen anzugleichen, die sie für erfolgreicher oder besser legitimiert halten. Eine solche isomorphe Haltung kann in Bezug auf die Ausrichtung der Ziele von Organisationen auf Shareholder Value durchaus beobachtet werden, wenngleich mit einigen Ausnahmen (Kochan & Rubinstein, 2000).

⁶³ Eine Orientierung an anderen Stakeholdern als den Anteilseignern kann für Unternehmen überdies, wie Harrison, Bosse & Phillips (2010) konzeptionell und empirisch zeigen, erhebliche Wertschöpfungsvorteile bringen, da diese Stakeholder so mehr angeregt werden in das Unternehmen zu investieren, etwa durch die Bereitstellung von Informationen. Im Gedankengebäude der Agency-Theorie dagegen wird dem Management empfohlen, den Stakeholdern exakt so viel Aufmerksamkeit zukommen zu lassen wie notwendig ist, um sie an einer negativen Beeinflussung der Wertschöpfung der Organisation zu hindern (vergleiche ebendort sowie obige Ausführungen zur Agency-Theorie).

konvergenter Validität im Zuge der beobachteten Heterogenität der verwendeten Operationalisierungen (vergleiche hierzu auch Abschnitt 4.2.1) tatsächlich gewährleistet ist oder sein kann. Skepsis hieran wird regelmäßig geäußert (etwa von Powell, 2003b). Hierauf wird im Folgenden weiter einzugehen sein.

Ein Problem liegt hierbei in eben jenen impliziten Vertragsbeziehungen und den sozialen und politischen Konstrukten der Organisation. Eine Vielzahl an Akteuren bildet durch diese Vertragsbeziehungen „das Unternehmen“. Sie stehen, annahmegemäß, diesem Unternehmen zwar wohlwollend gegenüber, streben aber dennoch danach, ihren eigenen Nutzen zu maximieren. Zur Verteilung auf die Stakeholder steht insbesondere die Wertschöpfung des Unternehmens zur Verfügung. Um diese wird im Wettstreit zwischen den Vertragsparteien gerungen, sie wird zwischen den Akteuren verteilt (Coff, 1999). Fraglich ist nun, wo die Performanceoperationalisierung in diesem Prozess ansetzt: Misst sie die Wertschöpfung der Organisation oder das Ergebnis eines Verteilungskampfes? Ist letzteres der Fall, verletzt die Leistungsoperationalisierung die Forderung nach konvergenter Validität – Konstrukte, die das selbe messen sollen, messen nun anderes, nämlich neben der Leistung auch die relative Machtposition einer Akteursgruppe im Kampf um die Verteilung der Wertschöpfung.

Methodisch ist dies insbesondere problematisch, da die relative Verhandlungsmacht verschiedener Stakeholdergruppen als unbeobachtete Variable die Beziehung zwischen abhängigen und unabhängigen Variablen systematisch verzerren kann. Im Folgenden analysiere ich dieses Problem in größerer Tiefe. Zunächst werden der Prozess betrieblicher Wertschöpfung und das Verständnis der Organisation als System von Einflussgruppen (Stakeholdern) weiter vertieft; die Appropriationstheorie bringt beides schließlich zusammen und diskutiert die Verteilung der Wertschöpfung auf Stakeholdergruppen. Schließlich gilt es, exemplarische Evidenz dafür zu sammeln, dass die Forderung nach konvergenter Validität im Rahmen der Leistungsoperationalisierung tatsächlich regelmäßig verletzt ist. Dies wäre nicht der Fall, wenn die unterschiedlichen Performanceoperationalisierungen stark korreliert wären.

3.3.2 Wertschöpfung und ihre Verteilung auf Stakeholder-Gruppen

Leistung ist, wie in Kapitel 2.2.1 erarbeitet, das kumulierte Ergebnis der Handlungen der Mitglieder einer Organisation. Im Rahmen des organisationalen Produktionsprozesses erzeugen Unternehmen aus Ausgangsstoffen Güter und Dienstleistungen, denen ein höherer Wert beigemessen wird als er den ursprünglich eingesetzten Inputfaktoren zu Eigen war. Die Differenz zwischen Absatz- und Vorleistungen ist die betriebliche Wertschöpfung. Die Kostenfunktion des Produzenten definiert für diesen einen minimalen Preis, zu dem er seine Güter noch verkaufen kann, ohne vom Markt verdrängt zu werden. Dieser Preis ermittelt sich aus den Kosten für Vorprodukte, für Arbeitsleistung, Betriebsstoffe, Gemeinkosten und einem kalkulatorischen Unternehmerlohn. Der Konsument weist in seiner individuellen Präferenzstruktur den Gütern einen spezifischen Nutzen zu, der mit einer für ihn maximalen Zahlungsbereitschaft einhergeht und eine monetäre oder zumindest quantifizierte Bewertung seiner Nutzenerwartung bei Kauf beziehungsweise Konsum der in Frage stehenden Güter oder Dienstleistungen darstellt. Eine Austauschbeziehung über den Markt kommt zustande, wenn die maximale Zahlungsbereitschaft des Konsumenten mindestens genauso groß ist wie der minimale Angebotspreis des Produzenten. Sind die Preise deckungsgleich, kann der Austausch zwar zu Stande kommen, doch entsteht weder dem Produzenten noch dem Konsumenten ein

Nutzen, beide Akteure stehen der Frage nach Tausch oder Nichttausch indifferent gegenüber. In der Regel wird die maximale Zahlungsbereitschaft des Konsumenten jedoch den Minimalpreis des Produzenten übersteigen. In diesem Fall bildet sich ein Marktpreis, der zwischen diesen beiden Extrempreisen liegt. In dieser Situation generiert nun mindestens eine der beiden Seiten Überrenten. Die Überrente des Konsumenten, die Konsumentenrente, ist die Differenz zwischen seiner maximalen Zahlungsbereitschaft und dem (niedrigeren) Marktpreis, multipliziert mit der Zahl der gehandelten Güter. Die Überrente des Produzenten, die Produzentenrente, entspricht dem Betrag, um den der Marktpreis die Kosten des Produzenten übersteigt, multipliziert mit der Zahl der gehandelten Güter. Der volkswirtschaftliche Wohlfahrtszuwachs ist die Summe dieser beiden Teilrenten. Da beide Parteien auf eine Optimierung ihres Individualnutzens hin orientiert sind, werden sie versuchen, den Marktpreis so zu ihren Gunsten zu verschieben, dass sie einen möglichst großen Teil der Gesamtrente abschöpfen können (Schöler, 2004). Diese klassische Modellierung der Mikroökonomie lässt sich graphisch folgendermaßen darstellen:

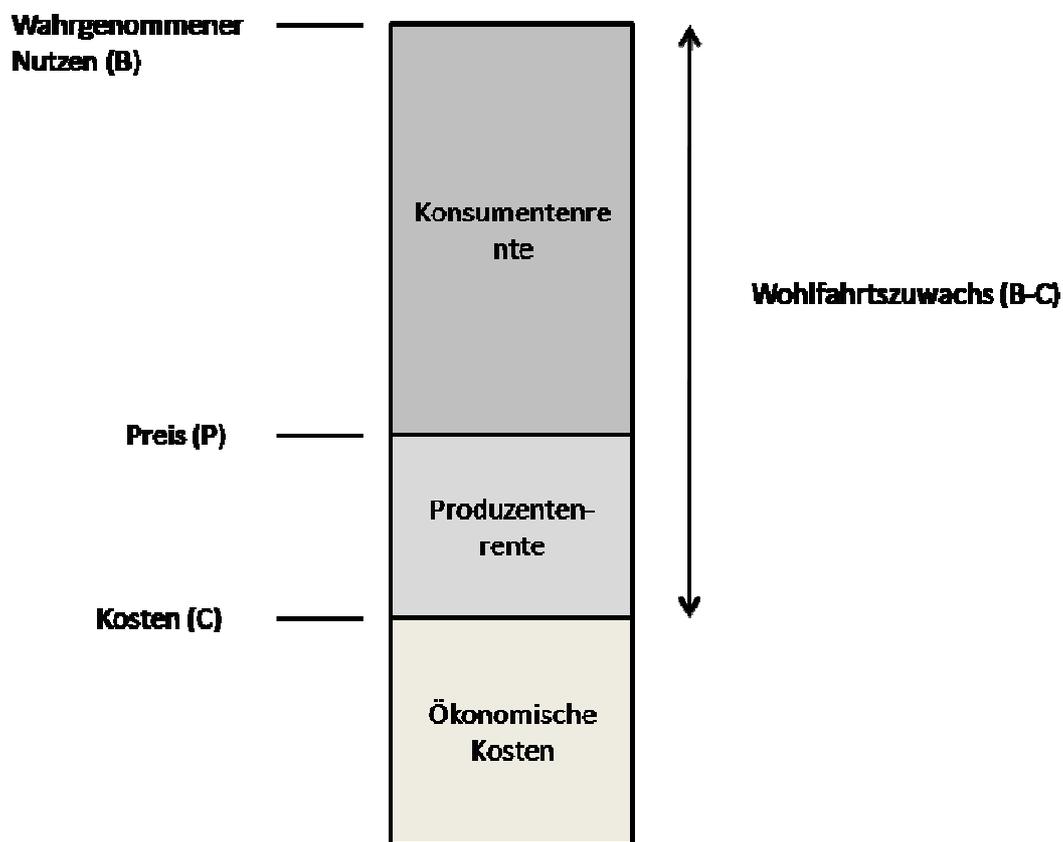


Abbildung 1: Ökonomische Rentengenerierung (eigene Darstellung, nach Peteraf & Barney, 2003, 314)

Die Herstellung der Güter verursacht ökonomische Kosten, den Faktorverbrauch (C)⁶⁴, und stiftet dem Käufer einen Nutzen (B), die Differenz (B-C) ist die Summe aus Produzenten- und Konsumentenrente, der gesamtökonomische Wohlfahrtszuwachs durch die Austauschbezie-

⁶⁴ Dieser Faktorverbrauch ist, wie erwähnt, nicht mit den betrieblichen Vorleistungen zu verwechseln, die im Rahmen der Definition von Wertschöpfung oben angesprochen wurden. Vorleistungen umfassen, anders als die hier dargestellten ökonomischen Kosten, nicht die Kosten für den Produktionsfaktor Arbeit, der im Sinne der weiteren Diskussion als Entgelt für geleisteten Input zuzüglich Appropriationsleistung der Arbeitnehmer verstanden werden kann.

hung. Welcher Anteil hiervon beim Produzenten verbleibt bestimmt der Preis (P): Je höher der Preis, desto höher der Anteil, den der Produzent vereinnahmen kann. Übersteigt der Preis den wahrgenommenen Nutzen des Käufers (B), kommt kein Austausch mehr zu Stande, sind beide deckungsgleich, ist der ökonomische Nutzengewinn gleich Null. Die vereinnahmte Produzentenrente (P-C) steht nun im Unternehmen wiederum als Verteilungsmasse für andere Stakeholder zur Verfügung. Auf die Produzentenrente versuchen durch implizite oder explizite Verträge gebundene, ihren Individualnutzen optimierende Personen und Gruppen Zugriff zu nehmen. Das Ergebnis dieses Verteilungsprozesses ist unternehmens- und situationsspezifisch (vergleiche Coff, 1999 und die Ausführungen weiter unten).⁶⁵

Akzeptiert man die oben vorgelegte Theorie der Organisation als Nexus impliziter und expliziter Verträge ist es notwendig, die Vertragsparteien einer genaueren Betrachtung zu unterziehen. Einen wesentlichen Beitrag zur Beschreibung des Verhaltens von Einflussgruppen in Organisationen liefert die Stakeholder-Theorie (vergleiche insbesondere die grundlegenden Beiträge hierzu: Freeman (1984) und Donaldson & Preston (1995)). Diese beschreibt die Organisation als System von Gruppen oder Einzelpersonen, die ein genuines Interesse an dieser Organisation haben. Diese Stakeholdergruppen sind grundsätzlich gleichrangig, es gibt also, im Gegensatz zum Neoinstitutionalismus, dem die Agency-Theorie zuzurechnen ist (Picot, Dietl & Franck, 2008), keine „natürliche“ Über- oder Unterordnung der Akteure. Die Stakeholder-Theorie setzt damit ein komplexes Bild des Wertschöpfungsprozesses im Unternehmen voraus. Im Umkehrschluss bedeutet dies nicht (notwendigerweise) die Ableitung einer normativen Aussage, etwa in der Form, dass ein Unternehmen nach der gleichmäßigen Erfüllung der Ziele aller Stakeholder streben sollte. Vielmehr beschreibt diese Feststellung den oben bereits konstatierten Zustand, dass zum Verständnis der Leistungserbringung einer Organisation, der deskriptiv-analytischen Dimension, die Ziele verschiedener Gruppen Beachtung finden müssen. Dem Management kommt eine koordinierende Rolle bei der Abwägung und Berücksichtigung der Ziele der unterschiedlichen Stakeholder zu.

Agency-Ansätze gehen dagegen von einem relativ simplen Input-Output-Modell aus. Demnach beliefern Zulieferer das Unternehmen mit Ausgangsstoffen, diese werden mit Hilfe von Arbeit und Kapital in Güter umgesetzt, die dem Kunden Nutzen stiften. Sowohl Zulieferer als auch Anbieter von Arbeit erhalten eine Normalvergütung ihres Inputs und agieren ansonsten im Gutenbergschen Sinne als Produktionsfaktoren (Gutenberg, 1973), sie sind also dispositiv und funktional zur Erstellung der Leistung (auch: Donaldson & Preston, 1995). Allenfalls treten sie, wie etwa in Porters „Five Forces“, als Gegner der Organisation in Erscheinung (Porter, 1999). Die Stakeholder-Theorie bricht mit diesem Modell und beschreibt ein System von Anspruchstellern an die Organisation sowie deren Wechselwirkungen untereinander. Mitarbeiter können demnach beispielsweise nicht mehr als dispositive Faktoren des Produktionspro-

⁶⁵ Diese simple Theorie makroökonomischer Wertschöpfung ist geeignet zur Analyse der dieser Arbeit zu Grunde liegenden Fragestellung und wird so auch in den hier zitierten Quellen, die den aktuellen Stand der Forschung darstellen, nicht weiter hinterfragt. Dennoch soll nicht unterschlagen werden, dass dieses Konzept eine Vereinfachung darstellt. Geschaffener Wert ist, wie auch der im Weiteren gelegentlich verwendete individuelle oder kollektive Nutzen, ein Gedankenkonstrukt, das empirisch nicht halt- oder quantifizierbar ist (Blankart, 2008). Für eine tiefere Diskussion des Konzeptes „Wert“ vergleiche Lepak, Smith & Taylor, 2007. Die Akzeptanz der einfachen Definition für diese Arbeit begründet sich rein instrumentalistisch daraus, dass nicht die Entstehung von Wert und Nutzen untersucht werden soll, sondern lediglich die forschungspraktische Rezeption seiner Verteilung. Hierfür sind die grundgelegten einfachen Theorien ausreichend präzise.

zesses mit einer Entlohnung nach Marktpreis verstanden werden. Vielmehr spielen sie, individuell oder organisiert, eine aktive Rolle im Prozess der Leistungserstellung. Sie versuchen für ihre Arbeitsleistung höhere Preise zu erzielen oder ihre Vorstellungen davon, wie das Unternehmen geführt werden sollte, in diesem durchzusetzen.⁶⁶ Ähnlich verhält es sich mit den anderen Stakeholder-Gruppen. Folgende Grafik veranschaulicht die zentralen Stakeholder in einem Unternehmen:

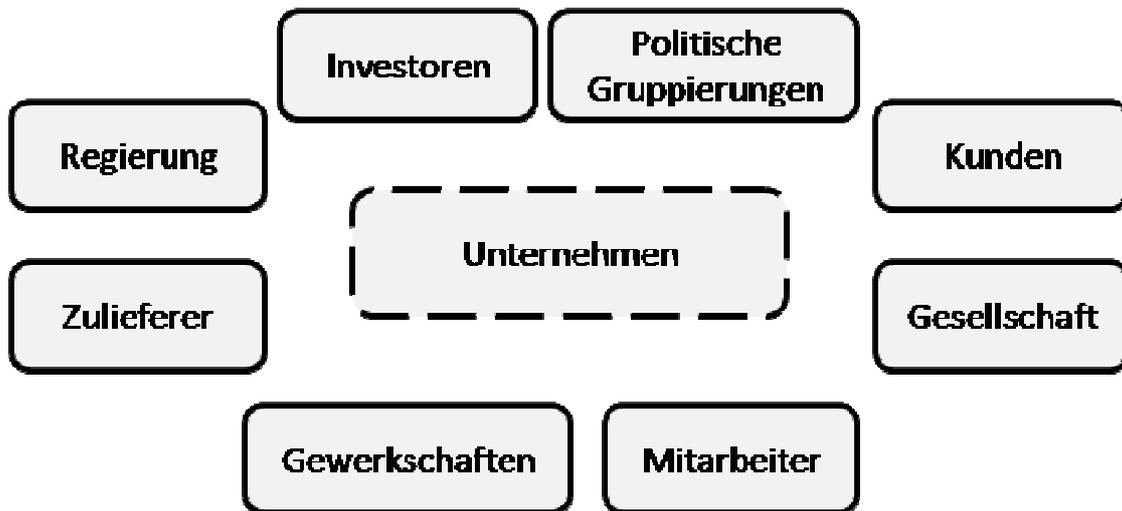


Abbildung 2: Stakeholder-Modell, (eigene Darstellung, angelehnt an Donaldson & Preston, 1995, 69)

Zu trennen ist damit zwischen dem Prozess der Leistungserstellung durch Produktion oder eine andere Schaffung von Werten und dem Prozess der Verteilung dieser Leistung auf all die Gruppen, die auf die erzeugten Werte Zugriff nehmen wollen (Lepak, Smith & Taylor, 2007). Für den Wissenschaftler stellt sich damit, neben der Kernfrage der Strategischen Managementforschung nach Ursprung und Beständigkeit von Wettbewerbsvorsprüngen auch die Frage, wie die geschaffenen Werte einer Organisation zwischen den Stakeholdern verteilt werden.⁶⁷ Die Verteilung der generierten Werte eines Unternehmens ist ein politischer Prozess, dessen Ausgang von der relativen Verhandlungsmacht der beteiligten Gruppen abhängt. Das Konzept der Bargaining Power, das diesen Betrachtungen zu Grunde liegt, findet insbesondere im Bereich Supply Chain Management und im Rahmen der Analyse von Allianzen und Netzwerkbeziehungen Verwendung, um die relative Verhandlungsmacht von Unternehmen und ihren Zulieferern oder Abnehmern beziehungsweise Netzwerkpartnern zu untersuchen (vergleiche etwa de Fontenay & Gans, 2007; Lippman & Rumelt, 2003; Crook & Combs, 2007). Auf intraorganisationale Prozesse angewandt liefert es jedoch auch einen starken Erklärungs-

⁶⁶ Vergleiche hierzu auch das von Denzau & North (1994) entwickelte Konzept der Shared Mental Models und dessen Anwendung auf die Stakeholder-Theorie durch Seisreiner (2006).

⁶⁷ Hierfür ist es auch keineswegs notwendig, in die bereits oben angesprochene normative Diskussion um die ethische Verpflichtung von Unternehmen einzusteigen (vergleiche hierzu Kacperczyk, 2009). Aus meiner Sicht ist die Begründung für Stakeholder-Theorien aus der oben diskutierten Kritik an der Shareholder-Orientierung ausreichend fundiert, ohne dass auf „fachfremde“, teilweise metaphysische Konzepte, etwa mit der Gerechtigkeitstheorie bei Donaldson & Preston (1995), zurückgegriffen werden müsste. Zur Problematik metaphysischer Ansätze in den Erklärungswissenschaften vergleiche Kapitel 2.1.

ansatz zur Analyse der Appropriation der in der Organisation geschaffenen Werte (vergleiche auch Hillman & Keim, 2001).

In seinem grundlegenden Beitrag zur Verteilung von generierten Renten in einer Organisation argumentiert Coff (1999) mit Ressourcen aus Sicht des Resource-based view. Demnach generiert ein Unternehmen Renten, wenn es über schwer imitierbare und nachhaltig wertschöpfungsrelevante Ressourcen verfügt und diese im Rahmen seines Leistungserstellungsprozesses einzusetzen weiß. Diese Ressourcen hängen beispielsweise mit intellektuellem oder prozeduralem Wissen zusammen, das in irgendeiner Form an Personen gebunden ist, etwa an die internen Stakeholder Management oder Mitarbeiter. Diese verfügen durch ihren Zugang zu diesen kritischen Ressourcen über Verhandlungsmacht gegenüber anderen Stakeholdern, wodurch sie in der Lage sind, einen gewissen Teil der in einem Unternehmen generierten Überrendite abzuschöpfen (Coff, 1999; Castanias & Helfat, 2001; Becerra, 2008). Auf eine ähnliche Art und Weise bringen auch andere Stakeholdergruppen spezifische Ressourcen in das Unternehmen ein, die zum einen wertschöpfungsrelevant sind, zum zweiten jedoch zu einer mehr oder weniger starken Machtposition der Stakeholdergruppe führen, mit der die Ressource verbunden ist (vergleiche Harrison, Bosse & Phillips, 2010).

Coff (1999) definiert vier Determinanten, anhand derer sich die relative Verhandlungsmacht einer Stakeholdergruppe bestimmt: Ihre Fähigkeit zu kollektiven Aktionen, ihr Zugang zu und ihre Kontrolle über kritische Informationen, die Kosten, die für das Unternehmen entstehen, wenn die Stakeholder das Unternehmen verlassen sollten sowie die Kosten, die für die Stakeholder selbst beim Verlassen des Unternehmens anfallen. Weitere Autoren ergänzten diese Systematik, etwa um Eigentumsrechte, die eine Stakeholdergruppe am intellektuellen Kapital einer Organisation ausüben kann (Skilton, 2009), institutionelle Regeln, die einer bestimmten Gruppe dienlich oder hinderlich sind (Chacar & Hesterly, 2008) oder die soziale Vernetztheit einer Stakeholder-Gruppe (Blyler & Coff, 2003).

Betrachtet man nun die Verhandlungsmacht der zentralen Stakeholdergruppen fällt auf, dass insbesondere die Anteilseigner im Vergleich zu Management und Mitarbeitern eine schwache Machtposition innehaben (vergleiche Tabelle 1). Problematisch ist, ähnlich wie auch in der Agency-Theorie diskutiert, der Zugang zu und die Kontrolle über Informationen. Auch ist das Kapital, das die Anteilseigner einbringen, nicht spezifisch und damit leicht zu ersetzen. Bei hohem Streubesitz ist auch die Fähigkeit zu konzertierten Handlungen der Anteilseigner stark beschränkt. Für die Verhandlungsmacht der Anteilseigner sprechen lediglich ihre sehr geringen Exit-Kosten.⁶⁸

Tabelle 1: Relative Verhandlungsmacht der internen Stakeholder (eigene Darstellung, angelehnt an Coff, 1999, 125)

Determinanten der relativen Verhandlungsmacht	Management	Mitarbeiter	Anteilseigner
Fähigkeit zu konzertierten Handlungen	+	?	-
Zugang zu und Kontrolle über kritische Informationen	+	+	-
Kosten bei Austritt des Stakeholders	+	+	-
Austrittskosten für Stakeholder	-	-	+

⁶⁸ Dies gilt zumindest im Fall größerer, börsennotierter Unternehmen, auf die sich die Arbeit von Coff (1999) implizit wesentlich bezieht.

Neben den von Coff diskutierten drei internen Stakeholdergruppen sind auch verschiedene externe Stakeholder darum bemüht, einen Anteil an der generierten Wertschöpfung zu vereinnahmen. Die Verhandlungsmacht externer Stakeholder basiert in der Regel weniger auf den Assets, die diese in das Unternehmen einbringen, als vielmehr auf dem Druck, den sie ausüben können. Marktkonzentrationen auf der Kunden- oder Lieferantenseite ist hierfür ein Beispiel (Asher et al., 2005). Die Analyse der relativen Verhandlungsmacht von Stakeholdergruppen und der organisationsinternen Prozesse, in denen Verteilungsergebnisse generiert werden, wird derzeit intensiv vorangetrieben, die Theorie der Verteilung der Wertschöpfung befindet sich dennoch im Moment eher am Anfang des wissenschaftlichen Verständnisprozesses (vergleiche die aktuellen Arbeiten von Skilton, 2009; Chacar & Hesterly, 2008; Jaros, 2008; Doucouliagos & Laroche, 2009; Coff, 2010). Die relative Machtposition von Managern gegenüber den Eigentümern ist hierbei sicher noch die am besten untersuchte Beziehung: In einer Untersuchung zur Appropriation von Überrenditen in F&E-intensiven Branchen durch Manager finden Coff & Lee (2003) einen signifikanten Zusammenhang zwischen Insider-Transaktionen durch Manager und Aktienkursentwicklungen der entsprechenden Unternehmen. Diese Insider-Deals können so als ein Vehikel gesehen werden, mit dem Manager Informationsasymmetrien, beziehungsweise Informationsvorsprünge gegenüber anderen Stakeholdern zur Appropriation von Überrenditen nutzen. Die anschließende positive Entwicklung des Aktienkurses deutet darauf hin, dass sich auch die Rentenappropriation der Anteilseigner fundamental positiv entwickelt, demnach wäre ein signifikant positiver Zusammenhang zwischen dem Insider-Trading von Managern und den angeeigneten Werten der Anteilseigner zu konstatieren. Zu beachten ist jedoch, dass hierdurch keinerlei Aussagen zum Anteil der Wertschöpfung möglich sind, der von den Anteilseignern appropriiert wird, angenommen kann nicht einmal zuverlässig auf eine Erhöhung der Wertschöpfung geschlossen werden. Es ist durchaus auch denkbar, dass sich lediglich die Informationsbasis der Anteilseigner und damit ihre relative Verhandlungsmacht erhöht. Kalyta (2009) untersucht opportunistisches Verhalten von Managern im Zusammenhang mit ihrer Möglichkeit, Gehaltsbestandteile, insbesondere Betriebsrenten, intransparent zu halten. Die Autoren finden, dass Manager in intransparenten Strukturen regelmäßig in der Lage sind, durch opportunistisches Verhalten größere Teile der Wertschöpfung zu vereinnahmen, und diese Möglichkeiten auch voll ausschöpfen, wenn ungenügende Transparenz die relative Machtposition der Eigentümer schwächt. Umgekehrt ist die Appropriation von Renten durch die Beeinflussung der eigenen Pensionspläne durch das Management weniger stark ausgeprägt, wenn den Eigentümern durch höhere Transparenz eine bessere Informationsbasis und damit mehr Verhandlungsmacht eingeräumt wird.

Klar wird hierbei, dass die Verteilung der generierten Renten in einem hoch komplexen politischen Prozess stattfindet, dessen Verlauf, zumindest nach aktuellem Wissensstand, von kau-saler Ambiguität geprägt und hochgradig dynamisch, interaktiv und situationsspezifisch ist (Frooman, 1999). Eine befriedigende theoretische Modellierung dieses Prozesses mit belastbarer empirischer Fundierung existiert derzeit meines Wissens nach nicht⁶⁹, doch liegen, ne-

⁶⁹ Einige empirische Studien liegen vor, diese untersuchen allerdings regelmäßig die Macht einzelner Stakeholdergruppen und die von Rowley (1997) kritisierten dyadischen Beziehungen zwischen der Stakeholdergruppe der Anteilseigner und dem Management, beispielsweise Murillo-Luna, Garcés-Ayerbe & Rivera-Torres (2008). Es fehlt ein umfassendes Verständnis der Interaktion der einzelnen Stakeholdergruppen.

ben der oben vorgestellten Modellierung von Coff (1999), einige Ansätze zur Erklärung der Verhandlungsmacht von Stakeholdern vor, wenn auch empirische Evidenz weitgehend fehlt. Mitchell, Agle & Wood (1997) ordnen Stakeholder nach drei Kriterien: Ihrer Macht, der Organisation zu schaden, der Dringlichkeit ihres Anliegens und der Legitimität desselbigen. Legitimität kann etwa auf juristischem Wege (explizite Verträge) oder auch auf einer moralisch-ethischen Basis gewonnen werden. Stakeholder, die in allen drei Kategorien punkten können, werden vom Management als am Relevantesten wahrgenommen, die Wahrscheinlichkeit, dass sie ihre Vorstellungen durchsetzen können, ist am höchsten. King (2008) beschreibt den Prozess der Organisation von Stakeholdern in Gruppen, eine Voraussetzung für effektive Einflussnahme in Organisationen. Der Einfluss unterschiedlicher Stakeholdergruppen hängt demnach wesentlich vom Fortschritt ihres gemeinsamen Meinungsbildungsprozesses, ihrer Aktivierung sowie äußeren Faktoren wie Schocks, denen die Organisation ausgesetzt ist, ab. In einer einfacheren Modellierung auf der Basis einer etablierten Theorie diskutieren Asher et al. (2005) die Einflussnahme von Stakeholdern aus Sicht der Theorie der Property Rights. Stakeholder halten demnach Anteile an Verfügungsrechten, die ihnen Zugriff auf die generierte Rente ermöglichen. Solche Property Rights können beispielsweise in der Form von Betriebsrenten, Tarifverträgen für Arbeitnehmer oder Entscheidungsbefugnissen und Aktienoptionen für Manager (Beispiele aus Asher et al., 2005, 8) vorliegen. Bezugspunkt dieser Property Rights sind die Ressourcen, die die Akteure in das Unternehmen einbringen. Je spezifischer und rarer diese Ressourcen, desto größer, beziehungsweise umfassender die Property Rights und damit auch die Verhandlungsmacht der jeweiligen Akteursgruppe. Die Argumentation von Asher und Kollegen ist damit sowohl nahe an der Austauschtheorie, die Herbert Simon bereits 1952 diskutierte (Simon, 1952), als auch am Ressourcenbasierten Ansatz, auf dem die Arbeit von Coff (1999) beruht. Bourne & Walker (2005) bewertet die Verhandlungsmacht von Stakeholdern auf der disaggregierten Ebene einzelner Personen nach der sich aus dem hierarchischen Status in der Organisation ergebenden Positionsmacht, der aus den eigenen Fähigkeiten und der sozialen Vernetzung resultierenden persönlichen Macht, sowie der politischen Macht, die etwa aus der Kontrolle von Entscheidungsprozessen, Koalitionen innerhalb der Organisation oder ähnlichem herrührt. Frooman (1999) dagegen sieht Macht nicht als Attribut von Personen oder Gruppen sondern von Beziehungen zwischen diesen Gruppen und wendet den Blick entsprechend auf diese Zwischenakteursebene; Rowley (1997) betrachtet die Beziehungen der Stakeholder als Netzwerk und kategorisiert die Reaktionen auf Druck von Stakeholdern durch die Organisation (vergleiche hierzu die Anmerkung unten) nach der Netzwerkdichte und der Zentralität eines Netzwerkakteurs.

Wie der derzeitige Forschungsstand zeigt, ist die Realität wenig mechanistisch und empirisch äußerst schwer fassbar. Die zentrale methodische Schwierigkeit bei der Messung der Rentenappropriation und der Verhandlungsmacht der Stakeholder liegt darin, dass nicht der Gewinn, sondern die Wertschöpfung, je nach Abgrenzung des Stakeholder-Feldes sogar noch Vorstufen hiervon, verteilt wird (Blyler & Coff, 2003; Becerra, 2008). Die Rentenappropriation vieler Stakeholder wird buchhalterisch über Kosten abgebildet, was die Ermittlung der konkreten Appropriationsleistung schwierig macht (Blyler & Coff, 2003). Verkomplizierend wirkt hierbei insbesondere auch die komplexe kausale Kopplung der Appropriationsleistungen (Becerra, 2008). Beispielhaft kann hierfür die Appropriation von Renten durch die externe Stakeholder-Gruppe „Gesellschaft“ sein, die sich unter anderem über CSR-Aufwendungen des Unternehmens abbilden lässt. Die Frage nach dem Zusammenhang zwischen finanzieller Per-

formance und CSR beschäftigt derzeit zahlreiche Forscher (etwa Godfrey, Merrill & Hansen, 2009; Barnett & Salomon, 2006; Hull & Rothenberg, 2008; Brammer & Millington, 2008). Aus Sicht der oben dargelegten Appropriationstheorie wäre zu erwarten, dass hohe CSR-Aufwendungen zu einer Reduzierung der finanziellen Performance führen, wenn diese als die Appropriationsleistung der Eigentümer (etwa in Form von CAR, TRS und ähnlichen Operationalisierungen) gemessen wird. Wie in Kapitel 3.3.3.2 noch näher ausgeführt wird, scheint diese Beziehung jedoch deutlich komplexer zu sein. So beobachtet Godfrey et al. (2009) etwa einen moderierenden Einfluss von CSR-Aufwendungen auf die Reaktion der Finanzmärkte im Falle negativer Ereignisse: Haben Unternehmen Aufwendungen für externe Stakeholder außerhalb ihrer unmittelbaren Aufgabenumwelt (Zulieferer, Kunden) getätigt, fällt die Kapitalmarktreaktion deutlich weniger negativ aus, als wenn dies nicht der Fall wäre. Ebenfalls relevant ist das Vertrauen, das zwischen den Stakeholdergruppen besteht und das in seiner Wirkung schwer fassbar ist. Verständnisleitend ist, wie oben ausgeführt, ein Menschenbild, das von bounded self interest, also nur beschränkt eigennutzmaximierendem Verhalten ausgeht. Denkbar ist, dass Maßnahmen, die auf den Aufbau von Vertrauen zwischen Stakeholdergruppen abzielen, zu einer Reduktion des eigennutzmaximierenden Verhaltens von Stakeholdern und damit zu einer Veränderung im Prozess der Wertverteilung führen können. Indizien für diese Annahme finden sich etwa in der Automobilindustrie (Pawlowsky & Wilkens, 1997). Forschungen zur Wirkung von Vertrauen auf Appropriationsprozesse in Unternehmen liegen derzeit jedoch noch nicht vor.

Allgemein ist die Analyse der Kausalketten im Bereich der Appropriation von Renten durch externe Stakeholder noch wenig durchleuchtet. Auch die Beziehung zwischen der Appropriationsleistung der Mitarbeiter und der der Eigentümer ist nicht weniger komplex (Choi & Wang, 2009). Die Appropriationstheorie, wie sie sich heute darstellt, kann einen kleinen Erklärungsbeitrag leisten, aber die komplexen Kausalitäten keinesfalls abdecken. Wesentliche Erkenntnis aus dem aktuellen Forschungsstand ist, dass die Summe der verteilten Wertschöpfung nicht gleich der Summe ist, die ursprünglich als gesamtökonomische Rente entstand. Vielmehr scheint durch die interaktive Kopplung der Stakeholdergruppen weiterer Wert generiert oder vernichtet werden zu können. Auf der anderen Seite kann die Appropriationstheorie aus dem bisher diskutierten heraus, anders als etwa Becerra (2008) anführt, nicht unmittelbar zur Erklärung der Unterschiede in der Wertschöpfung von Organisationen dienen. Zentral in diesem sich erst langsam zu einer Theorie entwickelnden Gedankengebäude ist die Verteilung der Wertschöpfung, nicht die Analyse ihres Zustandekommens. Wohl aber ist sie eng verwoben mit den Aspekten der Stakeholdertheorie, die positive Effekte auf die Wertschöpfung vermuten, wenn Stakeholder in besonderer Weise aktiviert werden können (Harrison et al., 2010). Es ist gut vorstellbar, dass Appropriationserfolg Stakeholder umgekehrt wieder dazu animiert, noch mehr in die Organisation einzubringen und damit die zur Verteilung zur Verfügung stehende Masse anwächst. Auf diese Weise ist dann auch wiederum ein Rückbezug auf die Wertschöpfung gegeben.

Interessant ist hierbei, dies sei an dieser Stelle als Exkurs noch eingefügt, die stete Verwendung der Dichotomie Stakeholder vs. Organisation, die sich, implizit oder explizit, insbesondere durch die Teile der Stakeholder-Forschung zieht, die sich nicht explizit mit der Appropriation von Renten befassen (etwa Frooman, 1999; Bosse, Phillips & Harrison, 2009; Brammer & Millington, 2008; Choi & Wang, 2009; Rowley, 1997). Versteht man die Stakeholder-Theorie wie oben dargelegt als Erklärungsmodell der Organisation, als Organisationstheorie,

die die Organisation als Nexus impliziter und expliziter Verträge und Ort der Beziehungen der Vertragspartner untereinander beschreibt, ist eine solche Unterteilung überflüssig bis irreführend. Die Organisation ist demnach durch Vertragsbeziehungen definiert und existiert nicht unabhängig von, sondern ausschließlich durch diese. Dies bedeutet, dass es keine eigenständige Entität „Organisation“ geben kann. Implizit wird, in oben erwähnter Dichotomie, unter Organisation entweder die Stakeholdergruppe Management oder, häufiger, die der Anteilseigner gefasst. Dies ist aus zwei Gründen problematisch. Zum Einen sind beide Parteien definitorisch ebenfalls Stakeholder (Freeman et al., 2004), ihre Benennung als „Organisation“ ist damit unangemessen und führt zu einer Wertung, beziehungsweise Priorisierung der Anteilseigner beziehungsweise des Managements, die aus der Theorie heraus nicht begründbar ist sondern ein Werturteil darstellt; zum Zweiten ist die unsaubere Trennung, beziehungsweise Vermengung von Anteilseigner und Manager im Begriff „Organisation“ dysfunktional, wie der gesamte Diskursraum der Agency-Theorie belegt (siehe oben). Diese konzeptionelle Unschärfe in der Verwendung des Organisationsbegriffs ist zur Analyse der Beziehungen von Stakeholdern untereinander daher als schädlich einzustufen.

3.3.3 Auswirkungen auf die Leistungsoperationalisierung: Verhandlungsmacht als Moderatorvariable

Die quantifizierte Leistungsmessung zielt häufig darauf ab, die vereinnahmten Werte einer bestimmten Stakeholder-Gruppe, typischerweise der Eigenkapitalgeber, zu erfassen. Die Konstruktion der Erfolgsvariablen entfernt sich so jedoch substantiell vom Kern des Unternehmenserfolges, der Generierung von Renten, und betrachtet weniger die eigentliche Leistung des Unternehmens als vielmehr das Ergebnis eines Verteilungskampfes. Würde man also die Messung der Leistung einer Organisation an Kennzahlen festmachen, die auf den Appropriationserfolg der Anteilseigner abstellen, würde man die tatsächliche Renditegenerierung einer Organisation systematisch falsch einschätzen, da eine wesentliche Variable, die relative Verhandlungsposition der Anteilseigner im Vergleich zu anderen Stakeholdern, ausgeblendet bleibt.⁷⁰ Um diese systematische Verzerrung zu vermeiden ist die Verwendung eines Erfolgsmaßes naheliegend, das den Verteilungsprozess so weit wie möglich aus der Erfassung eliminiert und sich auf diese Weise der Rentengenerierung einer Organisation annähert. In jedem Fall würde eine Leistungsoperationalisierung, die sich auf die Anteilseigner als einzig relevante Zielgruppe der Verteilung der generierten Renten bezieht, etwa auf dem Bilanzgewinn basierende Kennzahlen, den wesentlichen Teil der generierten Leistung systematisch ausblenden und stattdessen insbesondere die relative Verhandlungsmacht einer bestimmten Stakeholdergruppe messen. Das Bild der Leistung der Organisation wäre demnach verzerrt, die Forderung nach konvergenter Validität der Leistungsoperationalisierungen verletzt.

⁷⁰ Diese Argumentation steht im Konflikt etwa mit der Ansicht von Wolf (2008), der Shareholder- und Stakeholder-Orientierung als sich ausschließende Paradigmen im Kuhn'schen Sinne (vergleiche Kuhn, 2007) begreift. Diese Dichotomie scheint vor dem Hintergrund der hier diskutierten Erkenntnisse aus den Gebieten der Stakeholderforschung und der Methodik der Leistungsmessung aber unangemessen. Vielmehr kann die Shareholderorientierung als eine Sonderform der Stakeholder-Betrachtung verstanden werden, bei der die Anteilseigner eine institutionell stark legitimierte Position einnehmen (vergleiche zum Konzept der Legitimation Asher et al., 2005).

So merken auch Asher et al. (2005) an: „when assessing the economic value generated by the firm, one needs to assess the economic surplus captured by all stakeholders“ (S. 15). Markt- und accountingbasierte Kennzahlen dagegen messen nur den Teil der Wertschöpfung, der von den Anteilseignern vereinnahmt werden kann. Dies kann in Form des buchhalterischen Gewinns, beziehungsweise der ausgeschütteten Dividenden im Falle von Aktiengesellschaften oder aus diesem Gewinn abgeleitete Kennzahlen (accountingbasiert) geschehen oder indirekt über die Marktkapitalisierung, die als die Summe der zukünftig zu erwartenden abgezinsten Gewinne eines Unternehmens verstanden werden kann (Meyer, 2007), beziehungsweise mit aus der Marktkapitalisierung abgeleiteten Größen. In accounting- und marktbasierenden Kennzahlen ist also neben der Rentengenerierung der Organisation eine weitere Einflussgröße abgebildet, die relative Verhandlungsmacht der Anteilseigner im Vergleich zu der anderer Stakeholdern und ihre Fähigkeit, sich Anteile an der generierten Rente des Unternehmens anzueignen. Das wesentliche Problem im Rahmen der Performanceoperationalisierung liegt also in der Einbeziehung der Verhandlungsmacht als unerklärter Variablen, die den statistischen Zusammenhang zwischen unabhängigen Variablen und der zu erklärenden Zielgröße Leistung beeinflusst.

Wie im oben diskutierten Fall der Populationsselektion liegt das Problem auch hier nicht nur in der Einnahme einer bestimmten Sichtweise auf die Rentenappropriation (etwa der des Shareholder-Value Ansatzes oder einer Stakeholder-orientierten Perspektive) sondern auch in einer Suggestion von Aussagemacht, die möglicherweise bei bestimmten Performanceoperationalisierungen so nicht gegeben ist, wenn beispielsweise nur Teile der Leistung einer Organisation, etwa die durch die Eigentümer vereinnahmte, als Messgröße der Leistung fungieren. Auch hier gilt wieder, dass dies unproblematisch wird, wenn explizit die Appropriationsleistung einer bestimmten Gruppe untersucht werden soll. Dies ist zum Beispiel bei Berry (2006) der Fall, hier soll die Wertschätzung der Anteilseigner für bestimmte Managemententscheidungen gemessen werden. Marktbasierende Kennzahlen sind hier angebracht. Wohl aber muss die Leistungsspezifikation unter Nichtbeachtung der Appropriationsthematik als Problem gesehen werden, wenn etwa bei Kroll et al. (2007) unternehmensinterne Variablen und ihr Einfluss auf die Performance untersucht werden sollen, diese Performance aber an markt-basierten Kennzahlen gemessen wird. In diesem Fall wird eine Allgemeingültigkeit der Kausalität suggeriert, die so nicht vorliegen muss. Wie die Literaturreview in Kapitel 4 zeigen wird, liegt den meisten Arbeiten in der Strategischen Managementforschung ein breites Verständnis des Leistungsbegriffs zu Grunde, sie sind damit für das Problem der Appropriation als Moderatorvariable grundsätzlich anfällig. Die tatsächliche statistische Relevanz des Problems wird im Rahmen der Meta-Analyse in Kapitel 5 überprüft werden.

3.3.4 Die Appropriationstheorie und die unterschiedlichen Indikatoren

Das diskutierte Appropriationsproblem stellt sich forschungspraktisch nur, wenn unterschiedliche Leistungsoperationalisierungen tatsächlich unterschiedliches messen, also diskriminante Validität vorliegt, wo konvergente Validität erwartet, beziehungsweise beabsichtigt wurde. Sind die Operationalisierungen dagegen stark korreliert, stellt sich das Problem nicht. Einer ähnlichen Schwierigkeit sah sich die Innovationsforschung im Rahmen der Messung der Innovationsleistung gegenüber. Hier wird zumeist der gut zugängliche und leicht quantifizierbare Patentoutput als Operationalisierung der Innovationsleistung einer Organisation verwendet,

der jedoch aus verschiedenen konzeptionellen Überlegungen heraus diese nur unzureichend abbildet.⁷¹ Hagedoorn & Cloudt (2003) konnten jedoch zeigen, dass die Korrelation zwischen den üblichen Indikatoren zur Messung der Innovationsleistung (neben Patenten auch F&E-Input, Patent-Zitationen und die Einführung neuer Produkte) ausreichend hoch ist, um diese Indikatoren als gleichwertig anzusehen. Damit kann die Verwendung des am einfachsten handhabbaren Indikators Patentoutput, trotz seiner konzeptionellen Schwächen, forschungspraktisch gerechtfertigt werden.

Ähnliche Überlegungen müssen auch für das hier diskutierte Appropriationsproblem ange stellt werden. Wie dargelegt ist aus einer konzeptionellen Sicht heraus zu erwarten, dass die Appropriation des erwirtschafteten Mehrwertes durch Stakeholdergruppen und die relative Machtverteilung zwischen den Stakeholdergruppen in einer Organisation das Ergebnis einer Leistungsoperationalisierung verzerren. Dies ist forschungspraktisch jedoch nur dann von Relevanz, wenn unterschiedliche Indikatoren des gleichen Konstrukts (Performance) tatsächlich auch unterschiedliche Ergebnisse erbringen. Ist dies nicht der Fall ist es vollauf gerechtfertigt, einfach zugängliche Leistungsoperationalisierungen zu nutzen, etwa ROA oder TSR. An dieser Stelle soll der derzeitige Wissensstand zu solchen Korrelationen aus dem aktuellen Diskurs zur Leistungsoperationalisierung zusammengetragen werden. Die geschilderten Beispiele basieren auf Artikeln, die mit mehreren Leistungsindikatoren gleichzeitig arbeiten und dabei Korrelationskoeffizienten berichten. Im Folgenden werden dementsprechend explorativ die Zusammenhänge zwischen den wesentlichen Gruppen von Operationalisierungen⁷² und die Vorhersagen der Appropriationstheorie für erwartete Zusammenhänge zwischen diesen Indikatorengruppen diskutiert.

Accountingbasiert vs. marktbasier

Aus Sicht der hier vorgestellten Theorien sind zwei gegenläufige Tendenzen zu erwarten. Die Appropriationstheorie prognostiziert eine relativ hohe Korrelation zwischen accountingbasierten und marktbasieren Kennzahlen, da beide Indikatorengruppen die Aneignungsleistung der Eigentümer messen. Sowohl die Arbeitnehmerschaft als auch das Management haben zum Zeitpunkt der Erstellung der Bilanz oder der Rückflüsse für Aktionäre bereits auf die erzielte Wertschöpfung zugegriffen, auch Ausgaben im Bereich Corporate Social Responsibility (CSR), die als Proxy für die Appropriationsmacht externer Stakeholder gelten können (siehe unten), sind bereits wirksam geworden. Andererseits liegen unterschiedliche Zeithorizonte vor, die die Kausalität abschwächen könnten. Accountingbasierte Kennzahlen sind vergangenheitsorientiert, während sich in marktbasieren Kennzahlen die Erwartungen der Anteilseigner und der Teilnehmer am Private Equity Markt über zukünftig zu erwartende Zahlungsströme abbilden (vergleiche hierzu ausführlicher Uotila, Maula, Keil & Zahra, 2009; Zott & Amit, 2008; Hawawini et al., 2003; McNamara, Haleblan & Dykes, 2008).⁷³ Ebenfalls abschwächend auf die zu erwartende Korrelation sollte sich auswirken, dass accountingbasierte Kennzahlen Gegenstand bilanzpolitischer Gestaltungsmöglichkeiten des Managements sind,

⁷¹ Vergleiche den Beitrag von Hagedoorn & Cloudt (2003) auch für eine ausführlichere Diskussion der gebräuchlichen Innovationsindikatoren und ihrer Substituierbarkeit.

⁷² Zur Gruppierung von Indikatoren vergleiche Rowe & Morrow Jr. (1999) sowie die ausführlichere Diskussion in Abschnitt 4.1.3.2.

⁷³ Vergleiche zum Problem des Zeithorizonts auch Abschnitt 3.4.

während marktbasierende Kennzahlen diesem Problem weniger stark unterliegen (Hawawini et al., 2003). Anzunehmen wäre aus Sicht der Appropriationstheorie daher, dass accountingbasierte Kennzahlen die Aneignungsleistung der Stakeholder-Gruppe Management eher unterschätzen, da diese einen strategischen Anreiz hat, ihre relative Machtposition auf diese Weise nicht offen zu legen, da bei hoher Aneignungsleistung Gegenmaßnahmen der anderen Stakeholder-Gruppen wahrscheinlicher werden. Hierauf gibt Hawawini et al. (2003) einen Hinweis. Die Autoren untersuchen den Einfluss von Brancheneffekten auf die Performance und operationalisieren letztere über marktbasierende und accountingbasierte Kennzahlen. Während in einer unkorrigierten Stichprobe die Ergebnisse weitgehend identisch ausfallen, gehen diese deutlich auseinander, wenn die Stichprobe um Ausreißer, besonders gute und schlechte Unternehmen, bereinigt wird. Hier erklärt dann die Branchenzugehörigkeit einen großen Teil der marktbasierend gemessenen Performance, jedoch nur einen relativ kleinen Teil der accountingbasiert erfassten. David et al. (2008) nutzen ROA und Tobins q als Performance-Messgrößen und finden eine Korrelation von 0,4. Sanders & Hambrick (2007) dagegen untersuchen in einer Studie zum Risikoverhalten von Unternehmen Ausschläge von TSR und ROA und finden dabei nur eine äußerst geringe Korrelation von 0,05. Hillman, Shropshire & Canella Jr. (2007) nutzen ebenfalls Tobins q und ROA und ermitteln eine Korrelation von 0,32. Die Korrelation zwischen TSR und ROA liegt in derselben Studie bei 0,15. In einem ähnlichen Korrelationsbereich finden sich auch Market-to-Book Value und ROA in der Studie von Westphal & Stern (2007). Arthaud-Day, Certo, Dalton & Dalton (2006) finden eine Korrelation von 0,14 zwischen ROA und TSR und Choi & Wang (2009) eine von 0,35, wiederum zwischen Tobins q und ROA. Aus dem Rahmen fällt die Operationalisierung der Leistung mittels Cumulative Abnormal Returns (CAR), also den unerwarteten Änderungen der Marktkapitalisierung in einem bestimmten Zeitfenster. Hier kann praktisch keine Beziehung mehr zu accountingbasierten Kennzahlen festgestellt werden, McDonald, Westphal & Graebner (2008) etwa finden eine Korrelation von 0,01. Dies erklärt sich schlüssig aus den unterschiedlichen Betrachtungshorizonten: Accountingbasierte Kennzahlen decken eine gesamte Periode Geschäftstätigkeit ab, CAR dagegen messen spezifisch die Auswirkung eines isolierten Ereignisses. Zusammenfassend lässt sich festhalten, dass accounting- und marktbasierende Kennzahlen offenbar eine gewisse gemeinsame Basis aufweisen, zueinander jedoch nicht substitutiv zu sein scheinen. Die Argumentation der Appropriationstheorie wird von diesem Ergebnis unterstützt.

Accountingbasiert vs. CSR

Die Appropriation von Renten durch externe Stakeholder lässt sich sinnvoll über das Konzept der Corporate Social Responsibility (CSR) abbilden. CSR wird etwa von Hull & Rothenberg (2008) definiert als Maßnahmen, die von direktem Nutzen für die Gesellschaft, aber keinem oder allenfalls indirektem Nutzen für das Unternehmen selbst sind.⁷⁴ Mit „Unternehmen

⁷⁴ Diese Definition ist sicher nicht unstrittig, da von CSR-Aufwendungen Image-Effekte ausgehen, die wiederum einen unmittelbaren Einfluss etwa auf die Chancen des Unternehmens, Mitarbeiter zu akquirieren haben. Diese Mitarbeiter wiederum könnten die Performance steigern und so auch die Appropriationsleistung der Anteilseigner. Die Kausalitäten sind hier komplex und noch nicht vollständig erforscht, vergleiche hierzu auch Menguc & Barker, 2005, Godfrey, Merrill & Hansen, 2009 oder Barnett & Salomon, 2006. Für die Zwecke dieser Betrachtung bleibt jedoch festzuhalten, dass das CSR-

selbst“ ist im Kontext der Appropriationstheorie hauptsächlich die Stakeholder-Gruppe der Eigentümer gemeint, die „Gesellschaft“ lässt sich als Proxy für die Gesamtheit der organisati-onsexternen, nicht über Marktbeziehungen verbundenen Stakeholder verstehen. Die Unterscheidung zwischen CSR-Performance und accountingbasierter Performance bildet damit eine ausreichend trennscharfe Projektionsfläche, um eine entsprechende Korrelation untersuchen zu können. Aufgrund des relativ diffusen Charakters des CSR-Konzeptes ist die Untersuchung jedoch nicht einfach und die Ergebnisse, die in der Literatur zu finden sind, fallen entsprechend uneindeutig aus. In einer umfangreichen Meta-Studie finden Orlitzky et al. (2003) in der Tat eine positive Verbindung zwischen der finanziellen und der sozialen Performance von Unternehmen. Dies würde darauf hindeuten, dass das Problem der unterschiedlichen Stakeholder-Perspektiven in der Leistungsmessung von geringerer Relevanz ist und auf finanzielle oder soziale Leistung abstellende Messgrößen unter Umständen zueinander substitutiv wären. Allerdings verkompliziert sich die Analyse der Autoren, wenn für die spezifischen Operationalisierungen der sozialen Performance kontrolliert wird. Wird etwa die soziale Performance durch Reputations-Indices gemessen, ist die Korrelation stark, ein Halo-Problem wird hier vermutet, das heißt die bekannt gute finanzielle Performance der Unternehmen strahlt aus und lässt es auch in anderen Bereichen positiv erscheinen und umgekehrt. Betrachtet man dagegen die Korrelation zwischen finanzieller Performance und objektiv bewertbaren Prozessen der sozialen Performance (etwa die Entwicklung der von den betrachteten Unternehmen verursachten Umweltverschmutzung) oder Audits im sozialen Bereich (Bewertung der sozialen Aktivitäten durch neutrale dritte Parteien) wird der Zusammenhang deutlich schwächer. Ähnliches lässt sich auch für die Betrachtung der Operationalisierung der finanziellen Performance feststellen, so ist etwa der Zusammenhang zwischen von den betrachteten Unternehmen veröffentlichten CSR-Maßnahmen und accountingbasierten finanziellen Leistungsindikatoren (beispielsweise ROI) leicht negativ. In einer weiteren Meta-Analyse versuchen Heugens & Lander (2009) unter anderem, den Zusammenhang zwischen finanzieller Performance und der Befriedigung der Interessen externer Stakeholder zu ergründen. Letztere bezeichnen die Autoren als „symbolic performance: the extent, to which they generate positive social evaluations“ (S. 63); auch diese lässt sich nach obiger Definition unter dem Oberbegriff CSR fassen. Die unterstellte Kausalität ist, dass diese symbolische Performance den Unternehmen die Legitimation für die Aufrechterhaltung ihrer Tätigkeit verschafft. Die Autoren können einen positiven Zusammenhang bestätigen, allerdings ist die Erklärungsmacht der statistischen Modellierung gering.

In weiteren Studien wird eine positive Korrelation (und damit eine mögliche Substituierbarkeit der Kennzahlen) von verschiedenen Autoren bestritten, bzw. kann häufig nicht bestätigt werden. Nach den Erkenntnissen von Barnett (2007) etwa unterliegt der Zusammenhang zwischen der sozialen und finanziellen Performance eines Unternehmens einer starken Pfadabhängigkeit. Die Möglichkeit, „gute Taten“ in positive finanzielle Performance umzusetzen ist nach Barnett eine erlernbare Fähigkeit eines Unternehmens. Hull & Rothenberg (2008) betrachten die Frage nach dem Zusammenhang zwischen finanzieller und sozialer Performance aus einer industriestrukturorientierten Sichtweise und finden einen relativ starken Zusammenhang in eher homogenen Industrien, die sich durch eine geringe Produktdifferenzierung und wenig Innovationen auszeichnen. In dynamischen Industrien

Konzept die bestmögliche derzeit verfügbare Annäherung an die Appropriationsleistung der unternehmensexternen Stakeholder bietet und folglich auch hier Verwendung finden soll.

kann ein Zusammenhang dagegen nicht nachgewiesen werden. David, Bloom & Hillman (2007) stellen die Ausgaben für CSR in einen Erklärungskontext mit den Aktivitäten anderer Stakeholdergruppen und halten fest, dass Unternehmen, in denen die Anteilseigner aktiv für höhere CSR-Ausgaben plädieren, diese effektiv niedriger ausfallen als in Organisationen, in denen dies nicht der Fall ist. Die Autoren erklären dies damit, dass im Falle derartiger Forderungen von Seiten der Anteilseigner das Management eher dazu angeregt wird, Ressourcen zur Verteidigung der eigenen relativen Machtposition einzusetzen. Dieser Anteil an der Wertschöpfung stünde dann, im Betrachtungsrahmen der hier vorgestellten Appropriationstheorie, nicht mehr zur Verteilung zur Verfügung sondern würde, so die Autoren, wirkungslos in politischen Prozessen verbraucht. Zusammenfassend ist zu konstatieren, dass bisherige Untersuchungen nur in Einzelfällen starke Korrelationen nachweisen konnten, häufig ist der Zusammenhang zwischen CSR und accountingbasierten Leistungsindikatoren schwach ausgeprägt bis negativ. Die Vorhersagen der Appropriationstheorie werden also bestätigt.

Accountingbasiert und marktbasier vs. prozessbasiert

Unter „prozessbasiert“ werden alle Operationalisierungen subsumiert, die nicht auf der subjektiven Einschätzung von Experten, der Gewinn- und Verlustrechnung oder dem Marktpreis von Eigen- oder Fremdkapital beruhen. Viele prozessbasierte Größen wie etwa Überleben/Scheitern, Umsatz, Marktanteil oder Innovationsleistung stehen konzeptionell der Wertschöpfung nahe und messen weitgehend vor der Appropriation von Renten durch Stakeholder. Aus Sicht der Appropriationstheorie wäre daher eine schwach ausgeprägte Korrelation zwischen prozessbasierten und markt- und accountingbasierten Leistungsindikatoren zu erwarten. Nur sehr wenige Studien arbeiten sowohl mit accounting- als auch mit prozessbasierten Kennzahlen. Collis, Young & Goold (2007) untersuchen die Struktur von Corporate Headquarters. Hierzu befragten die Autoren Experten in den Headquarters nach ihrer Einschätzung zu Prozessgrößen (allgemeine Effizienz des Unternehmens und Kosteneffizienz) und erhoben gleichzeitig die finanzielle Performance mittels normalisierter ROCE sowie marktbasier als TSR. Die Ergebnisse bestätigen die Erwartungen der Appropriationstheorie. So finden die Autoren eine starke Korrelation zwischen TSR und ROCE (Korrelationskoeffizient 0,26) und eine noch stärkere zwischen den beiden prozessbasierten Größen (0,42), die Korrelation zwischen der Prozessgröße allgemeine Effizienz und der accountingbasierten Größe ROCE ist dagegen mit einem Korrelationskoeffizienten von 0,11 relativ schwach ausgeprägt. Nicht vollständig ins Bild fügt sich die Korrelation zwischen TSR und allgemeiner Effizienz. Hier messen die Autoren einen Korrelationskoeffizienten von 0,17, zu erwarten gewesen wäre ein Wert im Bereich der Korrelation zwischen Effizienz und ROCE von etwa 0,11. Von Nordenflycht (2007) arbeitet mit Umsatz pro Mitarbeiter und Profitabilität und findet keinen Zusammenhang zwischen diesen beiden Operationalisierungen. Profitabilität allerdings wird nicht näher spezifiziert, die konkrete Operationalisierung bleibt damit unklar. In derselben Arbeit findet der Autor einen starken Zusammenhang zwischen Wachstum und Überleben, zwei prozessbezogene Größen.

Accountingbasiert vs. Experteneinschätzungen

Aus Sicht der Appropriationstheorie lässt sich keine Aussage zu einer wahrscheinlichen Korrelation treffen, da das Leistungsverständnis des Experten nicht abschätzbar und die konkrete Form der Abfrage in der Studie häufig unklar ist (Wall et al., 2004). Studien, die eine Experten-Perspektive mit quantifizierbaren Einschätzungen aus dem Accountingbereich kombinieren sind selten, und Einschätzungen zur möglichen Korrelation entsprechend schwierig. Stam & Elfring (2008) schätzen die Leistung der von ihnen untersuchten Organisationen zum Einen mittels Befragung des Top-Managements ab, zum anderen nutzen sie Umsatzwachstum als accountingbasierte Kenngröße. Die Autoren finden eine recht starke Korrelation (0,32) zwischen diesen beiden Messgrößen. Als abhängige Variable in multiplen Regressionen verwendet ändern sich Koeffizientenwerte und Signifikanzniveaus in eher geringem Maße, die Richtung der Kausalität ändert sich nicht. Ähnlich stellt sich die Situation bei Morgan, Vorhies & Mason (2009) dar: Die Korrelation zwischen der subjektiven Einschätzung der Profitabilität durch interne Experten und der accountingbasierten Messgröße ROA beträgt in dieser Studie 0,32. Auf dieser Basis lässt sich eine vollständige Substituierbarkeit der Messgrößen zwar nicht bestätigen, doch scheinen Experten und accountingbasierte Kennzahlen die Leistung relativ ähnlich zu messen (vergleiche auch Wall et al., 2004). Dies wird im weiteren Gang der Arbeit detaillierter zu diskutieren sein.⁷⁵

3.3.5 Das Konzept der Kernperformance und ihre mathematische Modellierung

Analog der Betrachtung von Konsumenten- und Produzentenrente lässt sich auch für andere Stakeholder ein vergleichbares Verständnis von Überrendite zu Grunde legen. Im Fall der Gruppe der Mitarbeiter bedeutet dies beispielsweise, dass es einen (personenindividuellen) Minimalpreis geben würde, zu dem sie gerade noch in der Organisation arbeiten würden. Würde dieser Preis unterschritten, würden sie ihre Vertragsbeziehung mit dem Unternehmen beenden. Die Mitarbeiter werden im Rahmen des Vertrages danach streben, höhere Entlohnungen⁷⁶ zu realisieren. Die Überrendite, die die Mitarbeiter über ihren individuellen Mindestpreis hinaus, der auch als adäquate Kompensation ihres Inputs in die Organisation analog der Neoinstitutionalistischen Theorie verstanden werden kann (Becerra, 2008), akquirieren, stellt ihren persönlich vereinnahmten Anteil an der Gesamtrendite der Organisation dar. Die hierdurch verursachten direkten und indirekten Personalkosten schmälern unmittelbar den Anteil der Rente, der noch zur Verteilung auf andere Stakeholder zur Verfügung steht. Dabei ist ex post nicht mehr objektiv nachvollziehbar, welcher Teil der Personalkosten nun als Grundentgelt und welcher Teil als Überrendite verstanden werden kann, zumal dieser Anteil, wie angesprochen, personenindividuell ist und eine Aggregation über alle Mitarbeiter entsprechend komplex bis unmöglich erscheint.⁷⁷

⁷⁵ Vergleiche zum Einsatz von Experten in der Managementforschung auch Steigenberger, Pawlowsky, Lübcke & Schulte, 2010.

⁷⁶ Wie oben angesprochen umfasst Entlohnung hier die gesamte Gestaltung der Arbeit, nicht nur die pekuniäre Dimension des Arbeitsentgeltes.

⁷⁷ Vergleiche zum Konzept des Individualnutzens, das der Einfachheit und Klarheit der Argumentation halber hier Verwendung findet, volkswirtschaftliche Standardwerke, etwa Schöler (2004).

Modellhaft ist es möglich, aus einer Rückrechnung der von den jeweiligen Stakeholdern appropriierten Rentenanteile eine Kernperformance zu berechnen, eine Leistungsoperationalisierung vor allen Appropriationen. Mit Hilfe einer derartigen Modellierung kann ein vertieftes Verständnis des Leistungsverteilungsprozesses in der Organisation geschaffen werden.

Die Kernperformance, die auch als gesamtökonomische Rente verstanden werden kann, wird definiert als der Mehrwert, den alle beteiligten Stakeholder gemeinsam im Zuge des Wertschöpfungsprozesses einer Organisation über die jeweiligen Minimalpreise ihrer Inputs hinaus generieren. Diese gesamtökonomische Rente ist nicht mit dem gesamtökonomischen Wohlfahrtszuwachs identisch (vergleiche 3.3.2), da dieser sämtliche mit der Produktion einhergehenden Aufwendungen als Kostenblock vom resultierenden Gesamtnutzen subtrahiert. Dieses Verständnis ist nach der Appropriationstheorie zu einfach, hier sollen nur Kosten subtrahiert werden, die der Minimalvergütung der an der Erstellung des Gutes beteiligten Stakeholder entsprechen. Die gesamtökonomische Rente der Appropriationstheorie ist also größer als der gesamtökonomische Wohlfahrtszuwachs im Standardmodell.

Vom hypothetischen Konstrukt der Kernperformance ist zunächst der Anteil abzuziehen, den Zulieferer und Kunden als jeweils individuelle Produzenten- und Konsumentenrente im Rahmen der Belieferung der in Frage stehenden Organisation beziehungsweise dem Kauf der von dieser Organisation erstellten Güter und Dienstleistungen appropriieren. Fremdkapitalgeber können in dieser Modellierung als Lieferanten des Produktionsfaktors Kapital aufgefasst werden. Ihr individueller Minimalpreis bemisst sich nach der minimalen Verzinsung, die sie zu akzeptieren bereit wären. In der Kostenrechnung der Organisation schlagen sich überdies noch die Überrenten von Mitarbeitern und Managern nieder. Schließlich nimmt das Unternehmen Ausgaben für soziale Zwecke vor, bezahlt Mitgliedsbeiträge in Verbänden und so weiter, die Appropriation dieser externen Stakeholdergruppen. Vom Vorsteuergewinn ist schließlich noch die Steuer abzuziehen, die Appropriationsleistung der öffentlichen Hand. Auf den verbleibenden Gewinn greifen nun die Anteilseigner zurück, sowie über Erfolgsprämien gegebenenfalls erneut Mitarbeiter und Manager. Ein Teil der Wertschöpfung kann überdies als einbehaltener Gewinn wieder in die operative Arbeit des Unternehmens zurückfließen.

Mathematisch lässt sich die Kernperformance damit folgendermaßen modellieren:

$$(1) K = \sum_{i=1}^{\beta} (\sum_{j=1}^{\alpha} p_j^K * x_j^K) - \sum_{i=1}^{\gamma} (\sum_{j=1}^{\delta} p_j^Z * x_j^Z) - \sum_{i=1}^{\epsilon} p_i^A * x_i^A$$

Wobei p_j^K die maximale Zahlungsbereitschaft eines der β Kunden der Organisation für eines der α verkauften Güter des Unternehmens darstellt, x_j^K ist die von diesem Kunden gekaufte Menge des Gutes α . Das Unternehmen verkauft an β verschiedene Kunden, die äußere Summe addiert alle Preis-Mengen-Kombinationen aller Kunden. Der erste Term von Gleichung (1) stellt damit den hypothetisch möglichen Gesamtumsatz der Organisation bei vollständiger Abschöpfung der Konsumentenrente durch das Unternehmen dar. Hierbei handelt es sich um eine statische Betrachtung, das heißt die Kundenzahl, die Präferenz der Kunden, ihre jeweilige Zahlungsbereitschaft sowie das Güterportfolio werden als konstant angenommen. Gleiches gilt für die weiteren Betrachtungen, auch hier wird ein statisches Modell grundgelegt.

Der zweite Term in Gleichung (1) summiert in gleicher Art die Kosten für alle δ Güter aller γ Lieferanten, wenn die vollständige Produzentenrente der zugelieferten Güter oder Dienstleis-

tungen durch die Organisation, also den Kunden im Fall der Zulieferbeziehung, vereinnahmt worden wäre.

Im dritten Term schließlich wird der Produktionsfaktor Arbeit erfasst. p_i^A ist der Mindestpreis, für den ein bestimmter Mitarbeiter eine Einheit seiner Arbeitskraft der Organisation zur Verfügung stellen würde. x_i^A umfasst die Menge der vom Unternehmen in Anspruch genommenen Arbeitsleistung eines spezifischen Mitarbeiters. Diese Preis-Mengen-Kombination wird über alle ε Mitarbeiter aufaddiert. Es wird hier vereinfachend davon ausgegangen, dass die eingebrachten Arbeitsleistungen unterschiedlicher Mitarbeiter untereinander homogen sind. Das Management wird hier, ebenfalls vereinfachend, als Teil der Stakeholdergruppe Mitarbeiter aufgefasst.

Diese Kernperformance ist nun Gegenstand von Verteilungsprozessen. Wie in Abschnitt 3.3.2 geschildert, bildet sich bei Kauf und Verkauf von Gütern und Dienstleistungen in der Regel ein Marktpreis heraus, der zwischen dem Minimalpreis des Verkäufers und dem Maximalpreis des Käufers, der sich aus dessen mit monetären Größen bewerteter Nutzenerwartung bei Einsatz beziehungsweise Konsum dieses Gutes bestimmt, liegt. Sowohl Konsument als auch Produzent vereinnahmen einen Teil der generierten Rente. Entsprechend verhält es sich auch beim Arbeitslohn: In der Regel werden Mitarbeiter nicht zu ihrem individuellen Minimallohn arbeiten, sondern zu einem Lohn, der diesen übersteigt. Die Differenz ist ihre Überrendite, der Teil der Kernperformance, den sich anzueignen sie in der Lage sind.

Die Rentenappropriation der Kunden lässt sich entsprechend folgendermaßen modellieren:

$$(2) R_K = \sum_{i=1}^{\beta} (\sum_{j=1}^{\alpha} p_j^K * x_j^K) - \sum_{i=1}^{\beta} (\sum_{j=1}^{\alpha} p_j^{MK} * x_j^K)$$

p^{MK} ist hierbei der Marktpreis eines spezifischen an den Kunden verkauften Gutes. Entsprechend gilt für die Rentenappropriation der Zulieferer:

$$(3) R_Z = \sum_{i=1}^{\gamma} (\sum_{j=1}^{\delta} p_j^{MZ} * x_j^Z) - \sum_{i=1}^{\gamma} (\sum_{j=1}^{\delta} p_j^Z * x_j^Z)$$

Appropriiert wird die Differenz zwischen der mit Marktpreisen bewerteten über alle Lieferanten bezogenen Menge aller Güter abzüglich der mit Minimalpreisen bewerteten Preis-Güter-Kombination über alle Lieferanten. p^{MZ} stellt hier den Marktpreis der bezogenen Güter dar. Analog hierzu ist die Rentenappropriation der Mitarbeiter über Arbeitslöhne und ähnliche für das Unternehmen kostenwirksame Vergütungen der Arbeitskraft:

$$(4) R_A = \sum_{i=1}^{\varepsilon} p_i^{MA} * x_i^A - \sum_{i=1}^{\varepsilon} p_i^A * x_i^A$$

Auch hier ist p^{MA} wieder der Marktpreis einer Einheit, hier einer Einheit Arbeitsleistung. Der Marktpreis ist die tatsächlich an den Mitarbeiter ausbezahlte Vergütung. Die Differenz, die alle Mitarbeiter im Vergleich zu ihren jeweiligen Minimallohnen erzielen, ist in Summe die von den Mitarbeitern vereinnahmte Rente.

Aus (1) – (2) – (3) – (4) resultiert nun der in Abbildung 7 als „producer surplus“ firmierende Wert. Buchhalterisch lässt sich dieser Wert gut als Ergebnis der gewöhnlichen Geschäftstätigkeit annähern.

Wie Blyler & Coff (2003) anmerken findet sich die Rentenappropriation der Mitarbeiter schwer identifizierbar in den Kosten des Unternehmens wieder, (2) und (3) entziehen sich als möglicher entgangener Profit des Unternehmens beziehungsweise Produzentenrente der Zulieferer der wissenschaftlichen Erfassung ebenfalls weitgehend, da naturgemäß die individuellen Minimalpreise beziehungsweise -löhne nicht bekannt sind. Die weiteren Appropriationsschritte lassen sich dagegen relativ gut quantitativ fassen. Dieser verbliebene Teil der Kernperformance ist nun noch dem Zugriff weiterer externer Stakeholder wie gesellschaftlichen Gruppen, Interessensverbänden und so weiter ausgesetzt. Entsprechend ergibt sich für die nicht auf den Produktionsprozess bezogenen Appropriationen von Stakeholdern:

$$(5) R_E = \sum CSR + \sum Sonstige$$

Wie oben diskutiert können CSR-Aufwendungen als Proxy für die Appropriation gesellschaftlicher Gruppen aufgefasst werden. Der Sammelposten Sonstige steht für weitere Gruppen aus dem Umfeld des Unternehmens, wie eben Interessensverbände, Lobbygruppen oder ähnliche. Nachdem diese Gruppen auf den nach der Appropriation von Kunden, Zulieferern und Mitarbeitern noch verbliebenen Teil der Kernwertschöpfung zugegriffen haben, ergibt sich ein Vorsteuer-Gewinn, der entsprechend der betrieblichen Steuerbilanz noch der Appropriation von Werten von fiskalischer Seite ausgesetzt ist. Die Steuerlast lässt sich damit annähern durch:

$$(6) R_S = \{(1) - [(2) + (3) + (4) + (5)]\} * s;$$

$$\begin{aligned} R_S &= \{[\sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^K * x_j^K) - \sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^Z * x_j^Z) - (\sum_{i=1}^{\epsilon} p_i^A * x_i^A)] - [\sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^K * x_j^K) - \sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^{MK} * x_j^K)] - [\sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^{MZ} * x_j^Z) - \sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^Z * x_j^Z)] - [\sum_{i=1}^{\epsilon} p_i^{MA} * x_i^A - \sum_{i=1}^{\epsilon} p_i^A * x_i^A] - \sum CSR - \sum Sonstige\} * s = \\ &= \{\sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^{MK} * x_j^K) - \sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^{MZ} * x_j^Z) - \sum_{i=1}^{\epsilon} p_i^{MA} * x_i^A - \sum CSR - \sum Sonstige\} * s \end{aligned}$$

wobei s vereinfachend den rechnerischen Durchschnittssteuersatz des Unternehmens darstellt. Die Steuerbelastung, und damit die Rentenappropriation der öffentlichen Hand, bezieht sich damit also lediglich auf das Ergebnis nach der Appropriation von Renten durch Kunden, Zulieferer und Mitarbeiter. Schlussendlich appropriieren die Eigentümer noch einen Anteil d der verbliebenen Rente als Dividende oder Unternehmerlohn. d ist hierbei der Anteil des Nachsteuergewinns, der an die Eigenkapitalgeber ausgeschüttet wird.

$$(7) R_D = [(1) - (2) - (3) - (4) - (5) - (6)] * d;$$

$$\begin{aligned} R_D &= \{\sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^{MK} * x_j^K) - \sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^{MZ} * x_j^Z) - \sum_{i=1}^{\epsilon} p_i^{MA} * x_i^A - \sum CSR - \sum Sonstige - \\ &[\sum_{i=1}^{\beta}(\sum_{j=1}^{\alpha} p_j^{MK} * x_j^K) - \sum_{i=1}^{\gamma}(\sum_{j=1}^{\delta} p_j^{MZ} * x_j^Z) - \sum_{i=1}^{\epsilon} p_i^{MA} * x_i^A - \sum CSR - \sum Sonstige] * s\} * \\ &d \\ &= \text{Nachsteuergewinn} * d \end{aligned}$$

Der im Anschluss noch verbleibende Teil der Kernperformance bleibt als einbehaltener Gewinn R_U im Unternehmen.

$$(8) R_U = K - R_D - R_S - R_E - R_M - R_Z - R_K$$

Wie aus dieser mathematischen Modellierung ersichtlich ist, ist der Weg von der betrieblichen Rentengenerierung bis zu dem Teil, den die Anteilseigner appropriieren (R_D) weit. Die relative Verhandlungsmacht der Stakeholdergruppen beeinflusst alle mit einem M indexierten Preise sowie $\sum CSR$ und $\sum Sonstige$.

Diese Modellierung kann durchaus auch als Zeitachse verstanden werden. Auf die Kernwertschöpfung greifen zunächst die im Marktprozess befindlichen Akteure sowie die Bereitsteller von Produktionsfaktoren, zu, also Zulieferer, Kunden und Mitarbeiter. Diese Appropriation wird als Kosten beziehungsweise Umsatz buchhalterisch relevant. Auf den resultierenden Wert greifen dann Parteien zu, die nicht im Kernprozess des Unternehmens verankert sind, insbesondere externe Stakeholdergruppen. Schließlich bleibt ein buchhalterischer Vorsteuerertrag. Erst das Nachsteuerergebnis steht dann zur Verteilung auf die Anteilseigner zur Verfügung, diese können hiervon einen gewissen Teil in Form von Renditen auf eingebrachtes Eigenkapital (Dividenden, Unternehmerlohn etc.) vereinnahmen. Hierin muss nicht unbedingt eine Benachteiligung der Anteilseigner zu sehen sein, wie etwa Coff (1999) meint, vielmehr geht mit dem Residualcharakter des Anteils der Stakeholdergruppe der Eigentümer schlicht eine höhere Unsicherheit einher. Während die Appropriation der im Produktionsprozess gebundenen Stakeholder über die Zeit relativ konstant und vertraglich relativ gut explizierbar ist, bleibt den Anteilseignern (und auch der öffentlichen Hand) der „Restposten“, der sich wesentlich aus dem am Markt erzielbaren Umsatz bemisst. Dies kann je nach Geschäftsverlauf durchaus zu einer sehr hohen, aber auch zu einer geringen Appropriationsleistung der Anteilseigner führen. Sie tragen in der hier vorliegenden statischen Betrachtung also das unternehmerische Risiko, mit allen entsprechenden Vor- und Nachteilen.

Löst man die Annahme einer statischen Situation auf und lässt Variationen in den Marktpreisen und Input- und Outputmengen zu, erstreckt sich das unternehmerische Risiko jedoch auch auf alle anderen Stakeholder-Gruppen, da die Möglichkeit der Appropriation von der Fähigkeit des Unternehmens abhängt, entsprechende Verteilungsmasse zu schaffen, also eine hohe Kernperformance zu realisieren. Dieses Ergebnis kann durchaus als Untermauerung der Behauptung einiger Verfechter der Agency-Theorie gesehen werden, dass das, was den Aktionären nützt, den anderen Stakeholder ebenfalls nützt (Sundaram & Inkpen, 2004b). Allerdings ist der Satz umzuformulieren: Nicht eine Maximierung des Shareholder Value nützt allen Stakeholdern, sondern eine der Kernperformance. Von einer Beeinflussung der Marktpreise p^{Mx} profitieren in einer paretooptimalen Situation eben nicht alle Stakeholder. Diese Differenzierung unterbleibt bei Sundaram und Inkpen allerdings, ebenso wie bei anderen Verfechtern des Shareholder-Value Ansatzes.

Was bedeuten diese mathematischen Modellierungen nun für die Leistungsmessung in der Strategischen Managementforschung? Mit der Ableitung dieser Kernperformance verbindet sich zunächst keine normative Aussage hinsichtlich der Verwendung einer entsprechenden Performanceoperationalisierung. Sie stellt lediglich ein Konstrukt zur weiteren Analyse möglicher Verzerrungen durch Rentenappropriation dar. Wenn das Konzept der Kernperformance tragfähig ist, müsste die Verzerrung durch Appropriation stärker ausfallen, je weiter sich eine Performanceoperationalisierung von dieser Kernperformance entfernt. Ein relativ unverfälschtes Bild der Kernperformance sollten demnach Kennzahlen abbilden, die die Leistung möglichst früh im Verteilungsprozess messen. Hier bieten sich insbesondere Umsatzkennzahlen und auf den Produktionsprozess selbst fokussierte Kennzahlen wie die Innovationsfähig-

keit und das Überleben/Scheitern als letztgültiger Aufschluss darüber, ob ein Unternehmen in der Lage war, zumindest eine Kernperformance von 0 zu generieren, an. Eine Kernwertschöpfung von wenigstens 0 ist notwendig, um die Minimalpreise aller Stakeholder bezahlen und die Organisation so aufrechterhalten zu können. Auf der anderen Seite sind Kennzahlen, die die Appropriationsleistung einer bestimmten Stakeholdergruppe abbilden, insbesondere accounting- und marktbasierende Kennzahlen, wie die Modellierung zeigt, massiv vom Verteilungsprozess betroffen und bilden daher nur mehr zu einem geringen Teil die Kernperformance eines Unternehmens ab.

Fraglich ist jedoch, ob die Verwendung von Kennzahlen, die vom Verteilungsergebnis abstrahieren, überhaupt wünschenswert ist. Dies kann wiederum, ähnlich wie in Abschnitt 3.2 für das Themenfeld der Populationsselektion diskutiert, nur für den spezifischen Einzelfall entschieden werden. Wichtig ist die Aussagekraft ermittelter (oder auch nicht ermittelter) Kausalitäten rational einzuordnen, was im Fall der Kernperformance, die für sich ja wenig praktische Aussagekraft birgt, schwierig sein kann. Auch hier gilt es wiederum, Ziel und Mittel der Studie in Einklang zu bringen und im Rahmen der Arbeit das Problem der Rentenverteilung wahrzunehmen und in die konzeptionellen Überlegungen des Studiendesigns einzubeziehen. Zu klären ist nun, ob die Betrachtung unterschiedlicher Appropriationsstufen tatsächlich zu einer Verzerrung der Ergebnisse in der Strategischen Managementforschung führen kann. Lässt sich dies empirisch nachweisen, wäre zum einen eine starke empirische Unterstützung für die Appropriationstheorie gegeben, zum Zweiten ein mögliches Methodenproblem in der Strategischen Managementforschung im Rahmen der Ableitung kausaler Aussagen aufgezeigt. Die Meta-Analyse in Kapitel 5 wird diesen empirischen Test führen.

3.4 Das Time-Lag Problem

3.4.1 Die Problemlage, Hypothesenentwicklung

Ein Problem statistischer Untersuchungen im Allgemeinen und empirischer Arbeiten in der Strategischen Managementforschung im Besonderen liegt in der mathematisch schwer fassbaren Richtung von Kausalbeziehungen. Beeinflusst eine Variable X die Performance eines Unternehmens, oder ändert sich Variable X vielleicht nur, weil sich die Performance geändert hat? Die Untersuchung der Richtung von Ursache-Wirkungs-Zusammenhängen ist ein Grundproblem der mathematischen Statistik in der Ökonomie, da viele statistische Verfahren nur eine Beziehung, aber nicht die Richtung dieser Beziehung anzeigen können (Pearl, 2001; Bergh et al., 2004). Dieses Problem tritt beispielsweise bei Collis et al. (2007) auf. Die Autoren betrachten zur Erklärung ihrer „Variablen X“, der Struktur und Größe von Unternehmens-Headquartern, die Leistung der Organisation, gemittelt über einen 5-Jahres-Zeitraum vor den Untersuchungen. Diese Performance wird sodann in Relation zur Struktur und Größe der Headquarter gesetzt. Das Kausalitätenproblem hierbei ist klar: Die Autoren schlussfolgern (vereinfachend zusammengefasst) aus einem statistisch signifikanten negativen Zusammenhang zwischen relativer Größe des Headquartiers und der Performance, dass zu kleine Headquarter zu einer negativen Performance führen. Nicht entkräften können sie dabei den Verdacht, dass die Kausalität genau andersherum vorliegt: Die Unternehmen waren im untersuchten Zeitraum, dem Fünf-Jahres-Fenster vor Beginn der Erhebung, unterdurchschnittlich erfolgreich, haben sie möglicherweise deshalb ihre Headquarter verschlankt? Aus diesem

Untersuchungsdesign lässt sich keine klare Aussage zur Richtung der Kausalität treffen. Eine solche Aussage würde es erforderlich machen, die Performance in der Zukunft mit derzeitigen erklärenden Variablen in Zusammenhang zu bringen oder zumindest einen zeitlichen Versatz zwischen erklärender und erklärter Variablen anzunehmen (ähnlich beispielsweise auch in Sun et al., 2007). Dieses Problem umgekehrter Kausalitäten, oder des „confounding“ in der Terminologie von Pearl (2001, 173), deren Beitrag auch eine ausführliche mathematische Analyse liefert, stellt den Extremfall eines generellen Problems in der Strategischen Managementforschung dar: Des unklaren zeitlichen Zusammenhangs zwischen Ursache und Wirkung.

Die Generierung von Renten, das zentrale Erkenntnisinteresse der Strategischen Managementforschung, ist ein dynamischer und historischer Prozess (Schumpeter, 1952). Das Management eines Unternehmens beispielsweise trifft Entscheidungen, die Konsequenzen und Handlungen nach sich ziehen, die sich dann in einem bestimmten Leistungsergebnis niederschlagen. Äußere Umstände wie Umweltveränderungen, die Handlungen von Konkurrenten oder Präferenzveränderungen von Kunden (Porter, 1999) beeinflussen eine Organisation und verändern die Performance. Doch wie lange dauert es von einer Entscheidung des Managements bis zur Veränderung des Leistungsergebnisses? Welche zeitliche Verbindung besteht zwischen einer Umweltveränderung und einer bestimmten Performanceindikation? Allgemein: Welcher zeitliche und damit auch kausale Zusammenhang besteht zwischen den untersuchten Variablen?

Diese Schwierigkeit kann als Time-Lag Problem beziehungsweise Problem des Zeitversatzes ein wesentliches methodisches Problem in der Strategischen Managementforschung darstellen und damit zur Erklärung von Ergebnisheterogenität im Fachgebiet beitragen. Vom Moment einer Entscheidung bis zur Realisation der Konsequenzen in Form eines bestimmten Leistungsergebnisses vergeht Zeit (vergleiche auch McNamara et al., 2008). Je nach Fragestellung kann dieser Time-Lag eine Fehlerquelle oder zumindest ein Problem im Rahmen der Leistungsoperationalisierung in der Strategischen Managementforschung darstellen, nämlich dann, wenn zeitlich sensible Kausalitäten untersucht werden sollen. Henkel (2009) bezeichnet die in der Operationalisierung gemessenen Werte als „*empirical, or observed*“, die zeitlich eigentlich mit den erklärenden Variablen in einem kausalen Zusammenhang stehenden als „*true*“ (S. 288) und sieht im Auseinanderklaffen dieser beiden Werte ein zentrales Methodenproblem, in Fall seiner Arbeit für die Erfassung von Risiko.

Wenn die Leistung einer Organisation eine wesentliche Variable darstellt, stellt sich dieses Problem in besonderem Maße. Viele Operationalisierungen von Leistung weisen eine klare ex-post-Perspektive auf, sie wirken retrospektiv. Am eindeutigsten ist dies bei Größen, die sich auf die Kostenrechnung oder die externe Rechnungslegung eines Unternehmens beziehen, etwa ROA oder auch Umsatz- und Gewinngrößen. Hier wird explizit die Leistung einer vergangenen Periode erfasst.

Gleiches gilt auch, wenn die zu untersuchenden unabhängigen Variablen absehbar erst über einen längeren Zeitraum wirksam werden, wie es beispielsweise im Rahmen von Maßnahmen des Personalmanagements der Fall ist. Die Weiterqualifikation von Mitarbeitern, die Veränderung der Zusammensetzung von Managementteams, Maßnahmen, die die Motivation von Mitarbeitern beeinflussen sollen – all diese brauchen Zeit, um wirksam werden zu können. Eine Erfassung von abhängiger und unabhängiger Variablen in derselben Periode, hier etwa

die Einführung von Weiterbildungsprogrammen und die Performance eines Unternehmens, kann regelmäßig dazu führen, dass etwaige Zusammenhänge, die sich erst in der Zukunft zeigen werden, nicht aufgedeckt werden, etwa ein positiver Ergebniseffekt durch die bessere Qualifikation von Mitarbeitern in späteren Perioden. Kosten, die bereits in der Phase der Einführung anfallen, werden dagegen unmittelbar ergebniswirksam. So kann eine ungeeignete Spezifikation des zeitlichen Zusammenhangs zwischen Ursache und Wirkung dazu führen, dass in der Zukunft liegende Effekte verdeckt oder Maßnahmen, die sich erst zukünftig als positiv erweisen werden, mit einem negativen Leistungsergebnis assoziiert werden.

Erstaunlicherweise wird relativ selten für derartige Zeitverzögerungen kontrolliert oder die Notwendigkeit der Einbeziehung eines solchen Zeitversatzes zur Ableitung kausaler Aussagen gesehen⁷⁸, wie Greve & Goldeng (2004) in einer Literaturreview feststellen. In verschiedenen Studien könnte ein derartiger Time-Lag die gewonnenen Erkenntnisse zumindest stark relativieren. So untersuchen beispielsweise Thornhill & White (2007) den Einfluss einer archetypisch reinen Strategieausrichtung in Relation zu einer Strategie, die Charakteristika verschiedener generischer Strategietypen verbindet auf die Performance und kommen zu dem Ergebnis, dass sich Strategien, die sich eng an nur einem Archetypus orientieren, positiver auf die Performance auswirken als gemischte Strategien.⁷⁹ Die Strategie wird im Jahr 1999 ermittelt, der Erfolg anhand der Umsatzrendite des Jahres 2000 gemessen. Fraglich ist nun, inwieweit Strategien, die in 1999 implementiert und verfolgt wurden, bereits in 2000 umsatz- oder kostenrelevant werden können. Wiggins & Ruefli (2002) untersuchen in einer umfangreichen Zeitreihenanalyse die Dauerhaftigkeit beobachteter Überperformance. Die Autoren kommen zu dem Ergebnis, dass Überperformance nur sehr selten von längerer Dauer ist. Entsprechend wäre es entscheidend, den exakten Zeitpunkt der Realisierung der zu einer (beispielsweise strategischen) Entscheidung korrespondierenden Überperformance zu bestimmen, um überhaupt Kausalitäten finden zu können. Chang & Xu (2008) und Bergh et al. (2004) betonen die Gefahr umgekehrter Kausalitäten (reversed causalities), wenn nicht mit einem realistischen Zeitversatz in der Performance-Messung gearbeitet wird, Leistung und erklärende Variable also in der gleichen Periode erhoben werden. Wird ein solcher Zeitversatz einbezogen, erscheint zumindest die Gefahr umgekehrter Kausalitäten weitgehend gebannt, auch wenn damit die tatsächliche Kausalität noch nicht notwendigerweise nachgewiesen ist. Dieses Problem ist in der Literatur auch in einiger Tiefe diskutiert worden. Ein Überblick über diese Diskussion findet sich bei Bowen & Wiersema (1999), eine Analyse der Problemlage auch bei Devinney et al. (2005) und Greve & Goldeng (2004).

Ein weiteres Problem im Zusammenhang mit Zeiteffekten kann auch durch Zeitversatz im wissenschaftlichen Prozess selbst entstehen (Bergh et al., 2004), wenn zwischen Pre- und Post-Test oder zwischen der Messung der abhängigen und der unabhängigen Variablen (oder umgekehrt) so viel Zeit vergeht, dass sich Faktoren ändern können, die das Ergebnis der Untersuchung beeinflussen. Denkbar ist dies beispielsweise, wenn zuerst, etwa in einer Primärerhebung, die Leistung einer Organisation ermittelt wird, später dann erst deren Gründe erarbeitet werden, etwa in Fallstudien (so geschehen beispielsweise bei Peters & Waterman,

⁷⁸ Einige aktuelle Arbeiten mit expliziter Kontrolle für Zeitversatzprobleme: Hmieleski & Baron, 2009; Sharfman & Fernando, 2008 oder Chang & Xu, 2008, vergleiche auch Abschnitt 4.2.

⁷⁹ Die Autoren legen zwei Archetypen ihrer Untersuchung zugrunde: „Operating Excellence“, eine Strategie, die hauptsächlich auf Prozess- und Kostenoptimierung abzielt, und „Product Leadership“, hier steht die Produktentwicklung und Qualitätssicherung im Mittelpunkt.

1982). Allgemein tritt dieses Phänomen dann auf, wenn im Zuge der Erhebung unterschiedliche Messzeitpunkte relevant sind, zwischen denen sich die Umstände und Rahmenbedingungen geändert haben können.

Allerdings ist die bisherige empirische Evidenz für die forschungspraktische Relevanz solcher Time-Lags eher gemischt. Verschiedene Autoren kontrollieren für Time-Lags, ohne deutliche Änderungen an ihren Ergebnissen zu finden (etwa McDonald et al., 2008 oder Vroom & Gimenó, 2007). Wadhwa & Kotha (2006) allerdings zeigen, dass die in ihrem Artikel untersuchte Kausalverbindung zwischen der Aktivität als Venture-Capital Geber und dem Patent-Output sich nur bei einer Annahme eines Time-Lags von einem Jahr nachweisen lässt, ebenso wie auch Sharfman & Fernando (2008) in ihrer Suche nach einem kausalen Zusammenhang zwischen unternehmerischer Umweltpolitik und Kapitalkosten nur bei einem einjährigen Time-Lag fündig werden. Arbeiten, die konkrete Auswirkungen von Time-Lags auf Studienergebnisse auf einer breiteren Basis, etwa in Form einer Meta-Analyse untersuchen, liegen meines Wissens nach nicht vor, auch oben zitierte Literaturreviews bleiben auf einem deskriptiven Niveau. Weitere empirische Untersuchungen sind dringend geboten und werden in dieser Arbeit im Rahmen der Meta-Analyse in Kapitel 5 präsentiert.

3.4.2 Lösungsansätze

In der wissenschaftlichen Praxis werden verschiedene Möglichkeiten eingesetzt, um das Time-Lag Problem zu umgehen. Die wesentlichen Optionen stelle ich im Folgenden vor.

Die substanzreichste Möglichkeit zum Umgang mit Problemen des Zeitversatzes ist sicher die Bildung fundierter Hypothesen über die zeitlichen Beziehungen zwischen den zu untersuchenden Variablen. Solche Verfahren wurden in jüngerer Zeit beispielsweise von Lavie (2007), Wiersema & Bowen (2009a), Kacperczyk (2009), Acquaah (2007), Marcel (2009) oder Ellis, Reus & Lamont (2009) genutzt. Der übliche Versatz zwischen erklärender und erklärter Variablen beträgt hierbei eine Periode (beispielsweise bei Lavie, 2007 und Wiersema & Bowen, 2009a), während andere Autoren wie Acquaah (2007) oder Marcel (2009) auch mit drei- bis fünfjährigen Time-Lags rechnen. Die Dimension des Time-Lags muss hierbei naturgemäß von der zu untersuchenden Fragestellung abhängen. Notwendig hierfür ist die Verwendung von Panel-Datensätzen, da zeitversetzte Daten des gleichen Samples verfügbar sein müssen. Fiss & Zajac (2006), Richard et al. (2009) und Wall et al. (2004) betonen, dass dieses Vorgehen die Möglichkeit kausaler Aussagen deutlich erhöht und plädieren für die Verwendung langer Zeitfenster. Ihren Nutzen entfalten Panel-Datensätze allerdings nur, wenn auch für zeitverzögerte Performanceänderungen kontrolliert wird. Zudem weisen sie eine Reihe methodischer Schwierigkeiten auf, insbesondere in Form von Heteroskedastieproblemen sowie bei der Sample-Stabilität (Certo & Semadeni, 2006). Zapf et al. (1996) plädieren zur Bestätigung der Richtung von Kausalitäten für eine Zwei-Wellen Panel-Analyse: Nicht nur sollte der Einfluss der erklärenden Variablen zum Zeitpunkt t_0 auf die abhängige Variable zum Zeitpunkt t_1 geprüft werden, es sollte hierzu ergänzend auch betrachtet werden, ob sich ein Zusammenhang zwischen der abhängigen Variablen in t_0 und der unabhängigen Variablen in t_1 zeigen lässt. Ist dies der Fall, ist die Gefahr, dass eine umgekehrte Kausalität vorliegt, groß. Die Argumentation von Zapf et al. (1996) stammt aus dem medizinischen Bereich, der Untersuchung von Stress, scheint aber durchaus auch für Fragestellungen in der Strategischen Managementfor-

schung relevant. Hier sind solche Zwei-Wellen Panel-Analysen jedoch selten, eine der wenigen Anwendungen in der jüngeren Vergangenheit findet sich bei Siebert & Zubanov, 2009.

Eine weitere Möglichkeit um für das Vorliegen von Time-Lag-Effekten zumindest zu kontrollieren ist die Nutzung von zeitpunktbezogenen Dummy-Variablen, in der Regel Jahres-Dummies, für den Untersuchungszeitraum (etwa bei Wang & Zajac, 2007; Terlaak & King, 2007; Richard et al., 2007; Wasserman, 2006; Kim & Finkelstein, 2009 oder Oriani & Sobrero, 2008). Mit dieser Methode kann das jeweilige Jahr direkt in die Untersuchung mit einbezogen und auf statistische Signifikanz überprüft werden. Gute Ergebnisse sind hiervon zu erwarten, wenn die unabhängige Variable auf der Zeitachse fix ist und sich nicht ebenfalls über einen Betrachtungszeitraum erstreckt. Ist letzteres der Fall, kontrollieren Jahresdummies nicht mehr für Time-Lags sondern lediglich für systematische Veränderungen über den Zeitablauf, also für Beschränkungen, wie sie in Abschnitt 3.2 im Kontext der Populationsselektion diskutiert wurden. Eine unabhängige Variable ist insbesondere dann zeitlich fix, wenn es sich um ein bestimmtes Ereignis handelt, also etwa die Ernennung eines CEOs, ein Merger oder ähnliches. In einem solchen Fall lässt sich mit Jahresdummies gut prüfen, ob beobachtete Performance-Veränderungen mit einem bestimmten Zeitraum zusammenhängen.

Eine weitere Möglichkeit zum Umgang mit dem Time-Lag Problem in Zusammenhang mit Leistungsoperationalisierungen ist die Nutzung von Leistungsindikatoren, die selbst bereits über eine gewisse Prognosefähigkeit verfügen und damit, so die Hoffnung, zukünftige Entwicklungen bereits abbilden. Lebas & Euske (2007) regen zur Überbrückung dieses Time-Lags ein Instrumentarium an, „that will lead managers to taking appropriate actions in the present that will create a performing organization in the future“ (S. 127). Diese aus dem Blickwinkel des Business Performance Measurement und dem Gedankengebäude der Agency-Theorie formulierte Richtlinie kann durchaus auch als Empfehlung für die Überwindung der Zeitproblematiken in der Strategischen Managementforschung verstanden werden. Prognosefähigkeit wird insbesondere drei Kategorien von Performanceoperationalisierungen zugeschrieben: Marktbasierten Kennzahlen, innovationsbezogenen Indikatoren sowie Experteneinschätzungen. Innovationsbezogene Kenngrößen messen Produkt- oder Prozessinnovationen (oder Wege zu diesen), die zukünftig Erträge erwarten lassen und stellen damit einen mittelbaren Weg dar, zukünftige Wertschöpfung zu erfassen und so das Zeitversatzproblem zu umgehen. Die Hoffnung auf die Zukunftsbezogenheit marktbasierter Kennzahlen beruht auf der Annahme, dass sich im Marktwert die Erwartung der Anleger über eine zukünftige Wertentwicklung widerspiegeln und deren kollektive Intelligenz eine realistische Abschätzung zukünftiger Renditeerwartungen ermöglicht (Uotila et al., 2009; Zott & Amit, 2008; Devinney et al., 2005). Makri, Lane & Gomez-Mejia (2006) argumentieren in diesem Zusammenhang beispielsweise für die Verwendung von Market-to-Book Ratios zur Bewertung der Unternehmensstrategie: Wenn dieser Wert größer als 1 ist, ist das Unternehmen unterinvestiert, der Markt sieht mehr Wachstumspotential als sich in den Büchern wiederfindet. Umgekehrt wäre ein Wert von kleiner 1 so zu interpretieren, dass das Unternehmen zu stark investiert ist.

Allerdings scheinen an der Annahme, dass die kollektive Intelligenz der Marktteilnehmer die Möglichkeiten und zukünftigen Zahlungsströme von Unternehmen derart gut einschätzen kann, erhebliche Zweifel angebracht. Dies insbesondere vor dem Hintergrund der in Kapitel 3.3.1 und 3.3.2 diskutierten Probleme der Anteilseigner, Anteile der Wertschöpfung zu appropriateieren, die sich wesentlich eben auf eine schlechte Informationsversorgung dieser

Stakeholdergruppe zurückführen lässt. Der gesamte Argumentationsstrang der Agency-Theorie und damit des Shareholder-Value Paradigmas basieren auf der Annahme, dass Anteilseigner als Prinzipale über einen sehr schlechten Zugang zu Informationen verfügen und deshalb vor opportunistischem Verhalten anderer Stakeholder, insbesondere des Managements, zu schützen sind. Es muss aus einer konzeptionellen Sicht irritieren, dass diesen Prinzipalen nun eine derartig gute Einschätzung zukünftiger Erwartungen zugeschrieben wird. Auch die regelmäßig zu beobachtende hohe Volatilität in Unternehmensbewertungen muss vorsichtig stimmen. Levitas & McFayden (2009) betonen zudem, dass bei Verwendung marktbasierter Kennzahlen zur Abbildung von Zukunftserwartungen oder Risiko ein ausreichend langer Zeitversatz notwendig ist. Nur so ist gewährleistet, dass die relevanten Informationen zu den Marktteilnehmern diffundieren können und sich ein Marktpreis herausbilden kann, der tatsächlich Risiko und Zukunftserwartung abbildet. Ein solches langes Panel birgt allerdings für sich nun bereits wieder die Gefahr, dass die Kennzahl schlussendlich nicht mehr die Zukunft, sondern bereits wieder Gegenwart oder Vergangenheit abbildet und der spezifische Bonus marktbasierter Kennzahlen gegenüber etwa accountingbasierten wieder verloren geht. Auch das Vertrauen auf die kollektive Intelligenz der Anleger und die Fähigkeit der Aktienmärkte, die Werte von Unternehmen realistisch abzubilden, erscheint optimistisch. Mit den Problemen dieser Annahme befasst sich ein eigenes, wenn auch noch recht junges Fachgebiet, die Behavioral Finance, das in den Diskussionen zur Leistungsoperationalisierung in der Strategischen Managementforschung überraschenderweise bisher überhaupt noch nicht rezipiert worden ist (vergleiche etwa den Grundlagenartikel zur Behavioral Finance von Shiller, 2003). Die Annahme kollektiver Intelligenz von Marktteilnehmern ist jedenfalls vor diesem Hintergrund mit äußerster Skepsis zu bewerten.

Wie in Abschnitt 3.3 diskutiert, verfügen andere Stakeholdergruppen über einen deutlich besseren Zugang zu Informationen und bieten sich damit auch wesentlich eher zur Bewertung der Leistung eines Unternehmens an. Insbesondere die Befragung von Experten, oder key informants, wie sie im englischen Sprachraum gewöhnlich bezeichnet werden, lässt aus Sicht der Prognosefähigkeit der Leistungsoperationalisierung deutlich bessere Ergebnisse erwarten. Die Einbindung in Informationsströme reduziert die Agency-Problematik, die Verzerrung durch äußere Einflüsse, wie etwa in der Behavioral Finance diskutiert (etwa Spekulationsblasen) ist ebenfalls nicht in der Art gegeben. Auch die Objektivität von Experten ist durchaus in Zweifel zu ziehen, aus Sicht der Kognitionspsychologie etwa aufgrund der mannigfaltigen Einflüsse auf die Wahrnehmung von Ereignissen und die Beurteilung von Situationen durch Individuen (vergleiche beispielsweise Tversky & Kahneman, 1973; Wall et al., 2004; Ng, Westgren & Sonka, 2009). Auch ist zu hinterfragen, welche Anreize Key Informants haben, möglicherweise die Unwahrheit zu sagen, etwa, um ihre eigene Appropriationsleistung zu verschleiern und ihre Position im Verteilungskampf nicht zu verschlechtern. Doch scheinen diese Probleme deutlich weniger gravierend und durch die Auswahl geeigneter Informanten heilbarer als die Schwächen marktbasierter Kennzahlen (vergleiche allgemein zu Methodenproblemen beim Einsatz von Key Informants auch Lebas & Euske, 2007; Slater & Athuahene-Gima, 2004). Der Experte sollte als Insider am ehesten in der Lage sein, sich eine fundierte Meinung über zukünftige Entwicklungen zu bilden und diese zu kommunizieren. Die Korrektur für Time-Lag Probleme ist allerdings nur dann durch Experteneinschätzungen möglich, wenn diese auch zur erwarteten zukünftigen Entwicklung befragt werden. Werden lediglich Entwicklungen der Vergangenheit erfragt (wie beispielsweise bei Chandler & Hanks, 1994),

stellt die Befragung von Experten lediglich einen anderen Zugang zu accounting- oder prozessbasierten Kennzahlen dar.

Die einfachste Möglichkeit mit Problemen des Zeitversatzes umzugehen ist wohl die Verwendung von über einen Zeitraum aggregierten Performance-Größen. So addiert beispielsweise Luo (2008b) die ROIs seiner Untersuchungsobjekte schlicht über einen Zeitraum von drei Jahren auf. Hiermit wird das Problem zwar nicht gelöst, aber doch deutlich eingegrenzt. Diese simple Art der Überbrückung des Kausalitätenproblems findet sich in der Forschungspraxis allerdings selten.

Schlussendlich ist noch die Verwendung von Hazard-Modellen als Möglichkeit zum Umgang mit Zeitversatzproblemen zu nennen. Diese erfassen das Ausscheiden von Objekten aus einem Paneldatensatz über einen definierten Zeitraum, hier ist die Ermittlung des Zeitversatzes inkrementeller Bestandteil der Methodik, die somit gegenüber Zeitversatzproblemen robust wird. Die Verwendung von Hazard-Modellen bietet sich allerdings nur für ein sehr begrenztes Spektrum an Forschungsfragen an, da sichergestellt sein muss, dass das Ausscheiden aus dem Datenpanel tatsächlich als Ausprägung einer abhängigen Variablen interpretiert werden kann und nicht etwa aufgrund fehlender Daten erfolgte. Es ist damit im Kontext der Leistungsmessung lediglich dann praktikabel, wenn die Leistung eines Unternehmens als Überleben/Ausscheiden operationalisiert wird.

Ein Best Practice zur Überwindung oder zumindest Eingrenzung des Time-Lag Problems kann es hierbei kaum geben, vielmehr hängt die zu erwartende Zeitverzögerung zwischen Variablen maßgeblich vom zu untersuchenden Gegenstand ab. Entsprechend ist auch die Relevanz des Time-Lag Problems sehr unterschiedlich. Das Zeitversatzproblem zu ignorieren kann jedoch in vielen Fällen eine erhebliche Fehlerquelle darstellen und im Extremfall auch zum Problem der reversed causalities führen. Zumindest scheint die Hypothese gerechtfertigt, dass sich durch unterschiedlichen Umgang mit Time-Lag Problemen Teile der Heterogenität von Forschungsergebnissen in der Strategischen Managementforschung erklären lassen.

4 Qualitative Literaturreview

4.1 Forschungsansatz

4.1.1 Fragestellung

Nachdem in Kapitel 3 die konzeptionelle Ausleuchtung der in dieser Arbeit zu untersuchenden Problemfelder der Strategischen Managementforschung erfolgte, gilt es nun, einen ersten explorativen Überblick über den derzeitigen Umgang der Forschungspraxis mit diesen Problemfeldern zu gewinnen. Hierzu führe ich eine qualitative Literaturreview durch und beleuchte die Relevanz der Leistungsoperationalisierung in der Forschungspraxis, die Indikatoren, die im Rahmen der Operationalisierung des Konstrukts Leistung Verwendung finden sowie den Umgang mit den Problemen der Populationsselektion und der unklaren zeitlichen Zusammenhänge zwischen abhängigen und unabhängigen Variablen.

Hierzu untersucht die folgende Literaturreview Artikel, die in zwei Top-Journals des Themenkreises Management in der jüngsten Vergangenheit erschienen sind. Im Zentrum der Untersuchung stehen die folgenden vier Forschungsfragen:

Forschungsfrage 1 (F1): Wie relevant ist die Operationalisierung von Leistung für die Strategische Managementforschung?

Die in dieser Review betrachteten Journale stellen zentrale Medien zur Publikation qualitativ hochwertiger Arbeiten im Fachgebiet der Strategischen Managementforschung dar (vergleiche hierzu Abschnitt 4.1.3.1). Es kann daher angenommen werden, dass die in diesen Periodika publizierten Forschungsprojekte und -ergebnisse den derzeitigen wissenschaftlich-methodischen Stand zuverlässig abbilden. Von der Prüfung von Forschungsfrage 1 ist daher eine belastbare Aussage über die Relevanz der Leistungsoperationalisierung für die Strategische Managementforschung im Ganzen zu erwarten. Ziel ist, ganz im Sinne der qualitativen Review, ein erster Überblick über die Bedeutung der Leistungsmessung im Fachgebiet. Hierzu liegt bereits etwas ältere empirische Evidenz vor. March & Sutton (1997) untersuchten in einer ähnlichen Betrachtung das Academy of Management Journals (AMJ), das Strategic Management Journal (SMJ) sowie das Administrative Science Quarterly und analysierten dabei 427 Artikel, die zwischen 1993 und 1995 erschienen waren. Die Autoren finden für rund 30 Prozent der untersuchten Artikel eine Performanceoperationalisierung als unabhängige Variable, 20 Prozent als abhängige Variable. Auch Richard et al. (2009) führten vor kurzem eine ähnliche Untersuchung durch und finden für den Zeitraum 2005-2007 vergleichbare, wenn auch etwas geringere Werte.

Forschungsfrage 2 (F2): Findet sich ein dominantes Paradigma der Leistungsmessung in der Strategischen Managementforschung?

Die zweite Fragestellung, die die folgende Literaturreview beantworten soll, betrifft die Bandbreite möglicher Lösungsvorschläge zur Operationalisierung von Leistung. Wie erwähnt drückten verschiedene Autoren ihre Unzufriedenheit mit dem bestehenden Instrumentenkatalog zur Messung der Leistung von Organisationen aus, unter anderem Boyd et al. (2005), Chakravarthy (1986), March & Sutton (1997) oder Ray, Barney & Muhanna (2004). Aufgrund der Beobachtung, dass es keine allgemein akzeptierte Leistungsoperationalisierung zum Zeit-

punkt der Erstellung dieser Beiträge gab, ist auch in der aktuellen Forschung eine wesentliche Methodenpluralität zu erwarten. Zu überprüfen ist also, ob die Beobachtung des Fehlens eines dominanten Mess-Paradigmas auch im Betrachtungsraum der vorliegenden Literaturreview, dem Zeitfenster 2006 bis 2009, noch gilt. Diese Frage ist für die Argumentation der Arbeit auch von zentraler Bedeutung, da ihre Beantwortung die in Abschnitt 3.3.1 aufgegriffene Kontroverse um Annahme oder Ablehnung der Agency-Perspektive in der Forschungspraxis tangiert.

Würde Forschungsfrage 2 bejaht, könnte sich die weitere Arbeit auf diese dominante Operationalisierung konzentrieren. Wenn eine dominante Operationalisierung den Blickwinkel des Shareholder-Value Ansatzes einnehmen würde, könnte gezeigt werden, dass das Paradigma der Agency-Theorie für die Strategische Managementforschung derzeit verständnisleitend ist. Überdies lässt sich auch ein Überblick über das verwendete Methodenspektrum im Rahmen der Leistungsoperationalisierung in der Strategischen Managementforschung geben.

Im Rahmen der Forschungsfragen 3 und 4 gilt es dann, den Umgang mit dem Problem der Verallgemeinerbarkeit von Ergebnissen über Teilpopulationen sowie der unklaren zeitlichen Zusammenhänge zwischen abhängigen und unabhängigen Variablen zu beleuchten, wie sie in den Abschnitten 3.2 und 3.4 erarbeitet wurden. Ziel ist hier jeweils die Beschreibung des status quo im Umgang mit diesen beiden Problemstellungen.⁸⁰

Forschungsfrage 3 (F3): Nach welchen Kriterien werden Teilpopulationen abgegrenzt? Wird das Problem der begrenzten Aussagekraft, die durch die Verwendung begrenzter Teilpopulationen zu Stande kommt, in Arbeiten der Strategischen Managementforschung erkannt und benannt?

Forschungsfrage 4 (F4): Wird das Problem des Zeitversatzes erkannt? Welche Verfahren werden angewandt, um es einzugrenzen?

4.1.2 Qualitative Literaturreview – Methodik

Die folgende Erhebung nutzt die Technik der systematischen qualitativen Literaturreview (Green et al., 2006). Eine qualitative Literaturreview zeichnet sich allgemein dadurch aus, dass sie einen möglichst vollständigen Überblick über bisher zu einem Thema vorliegende Beiträge liefert, ohne die diesen Artikeln zu Grunde liegende Datenbasis weiter zu verwenden.⁸¹ Die betrachteten Artikel werden hierbei Analysen, Einschätzungen und Bewertungen unterzogen. Die systematische Variante erfordert, in Abgrenzung zur unsystematischen qualitativen Review, eine Reihe methodischer Fundierungen: So muss die Methodik der Auswahl und Bewertung der Artikel offen gelegt und nachvollziehbar sein, alle einbezogenen Beiträge müssen den selben Bewertungsmaßstäben unterworfen sein; zur Minimierung der Biases in der Analyse der Artikel wird die Review für gewöhnlich von mehreren Wissenschaftlern unabhängig und parallel durchgeführt, im Anschluss daran werden die Einschätzungen abgegli-

⁸⁰ In Anhang A.3.1 finden sich überdies eine Auswertung zur geographischen Verteilung von Arbeiten in der Strategischen Managementforschung, die nicht zum Argumentationsstrang dieser Arbeit gehört aber dennoch im Kontext einer Diskussion des Fachgebietes der Strategischen Managementforschung von Interesse ist.

⁸¹ In Abgrenzung zur quantitativen systemischen Literaturreview oder Metaanalyse (vgl. Green, Johnson & Adams, 2006). Eine solche ist die methodische Grundlage von Kapitel 5 dieser Arbeit.

chen und Zweifelsfälle diskutiert. Die Ergebnisse werden systematisch, zumeist in Form von Ergebnistabellen, präsentiert und kritisch diskutiert (Green et al., 2006, 104). Der qualitative Charakter dieser Methodik rührt von der Analyse von Einzelfällen her, die im Rahmen der Literaturreview nur erfolgen kann, es wird also nicht versucht, aus der gezogenen Stichprobe auf eine größere Grundgesamtheit zu schließen, methodisch beschränken sich alle Aussagen auf die gezogene Stichprobe. Der Anreiz zur qualitativen Literaturreview ist darin zu sehen, dass sie einen belastbaren Überblick über ein spezifisches Themenfeld ermöglicht, ohne Homogenität im Erkenntnisinteresse der betrachteten Artikel zu fordern, wie dies im Falle der quantitativen Literaturreview der Fall wäre. Sie eignet sich damit insbesondere zur Darstellung der Geschichte eines Problemfeldes, der Breite von Lösungsansätzen, zur Untersuchung einer Entwicklung sowie bei schlecht spezifizierten Fragestellungen in der Phase wissenschaftlicher Theorienbildung (Green et al., 2006; Barr, 2004). Die wesentliche Schwäche im Vergleich zur quantitativen Literaturreview liegt im Verzicht auf Informationen, die durch die empirische Weiternutzung von in empirischen Primärstudien gewonnenen Daten, die die Basis der quantitativen Literaturreview darstellt (vgl. Abschnitt 5.2), erlangt werden könnten.

In Abwägung der Stärken und Schwächen scheint die Methodik hier zielführend. Die vorliegende Review zielt auf eine Exploration des Feldes, auf einen Überblick über die in der wissenschaftlichen Praxis derzeit verwendeten Methoden. Dies erfordert eine Einbeziehung sehr heterogener Artikelgruppen und Diskurse, die mit einem quantitativen Ansatz nicht möglich wäre. Daher erscheint die qualitative Literaturreview in ihrer systematischen Ausprägung als das geeignete Mittel um die aufgeworfenen Forschungsfragen zu beantworten. Rynes, Giluk & Brown (2007) empfehlen dieses Vorgehen für Studien, die einen Überblick zum aktuellen Stand der Forschung zum Ziel haben und untersuchen selbst in ihrer Studie zum Praktiker-Wissenschaftler-Gap in der Human Resource-Forschung drei Praktiker-Journals in einer Vollerhebung. Die Methode der systematischen qualitativen Literaturreview wird in der Strategischen Managementforschung beispielsweise auch von Hoetker (2007) verwendet, in seiner Analyse der Verwendung von Logit- und Probit-Modellen, sowie von March & Sutton (1997) in einer Analyse zur Verwendung von Performanceoperationalisierungen. Dieses Verfahren scheint daher sowohl allgemein als auch spezifisch im in dieser Arbeit interessierenden Fachgebiet der Managementforschung geeignet, einen aktuellen Wissens- und Methodikstand systematisch aufzubereiten. Im Rahmen der Analyse zur Methodenverwendung wird empfohlen, möglichst aktuelle Forschungsergebnisse zu betrachten, da bei diesen die höchsten methodischen Standards zu erwarten sind (Boyd et al., 2005). Dem kommt der vorliegende Ansatz nach.

Die Auswertung dieses ersten empirischen Teils der Arbeit erfolgte, wie auch alle weiteren empirischen Analysen, mit Hilfe des Statistik-Software SPSS in der Programmversion 17.

4.1.3 Datenquellen und Operationalisierung

4.1.3.1 Die untersuchten Journals

Die Literaturreview basiert auf einer Vollerhebung zweier akademischer Journale über jeweils drei Jahrgänge, die sich thematisch mit dem Forschungsbereich des Strategischen Managements befassen und in verschiedenen Studien als besonders bedeutend für die wissenschaft-

liche Gemeinschaft identifiziert wurden.⁸² Die Beschränkung auf Top-Journals begründet sich durch den intendierten Fokus auf die Spitzenforschung und wurde in ähnlicher Form beispielsweise auch von Heugens & Lander (2009) empfohlen und durchgeführt. Hoetker (2007) greift in seiner Methoden-Review der Strategischen Managementforschung ebenfalls auf das Strategic Management Journal und das Academy of Management Journal zurück. Auch Boyd et al. (2005), Williams, Gavin & Hartman (2004) und Ketchen et al. (2008) empfehlen diese Journals für Literaturreviews in der Strategischen Managementforschung.

Im Einzelnen wurden betrachtet:

1. Das Strategic Management Journal (SMJ): Die Zeitschrift der Strategic Management Society erscheint mit 13 Ausgaben jährlich, die jeweils zumeist 5 oder 6 einzelne wissenschaftliche Beiträge umfassen. Das SMJ wurde 1980 ins Leben gerufen. Das Mission Statement des SMJ lautet:

*The SMJ promotes the development and dissemination of advances in the field [strategic management] by maintaining the highest standards of intellectual thought combined with practical relevance.*⁸³

Das Strategic Management Journal verfolgt eine starke (wenn auch nicht ausschließliche) empirische Ausrichtung und einen klaren Fokus auf das Strategische Management und eignet sich damit besonders zur Beantwortung der in dieser Literaturreview aufgeworfenen Fragestellungen (vergleiche auch Barney & Mackey, 2005). Zudem begleitet es das Fachgebiet der Strategischen Managementforschung maßgeblich seit einem sehr frühen Entwicklungsstadium dieser Diskurstadtion (Nerur, Rasheed & Natarajan, 2008). Das Strategic Management Journal ist meines Wissens nach das einzige internationale „A“-Journal, das sich exklusiv der Strategischen Managementforschung verschrieben hat.

Die Literaturreview beginnt mit Heftnummer 9, Jahrgang 27 (September 2006) und endet mit Heftnummer 8, Jahrgang 30 (August 2009).⁸⁴ Nicht in der Betrachtung enthalten ist Heft 13, Jahrgang 2006, das keine wissenschaftlichen Beiträge sondern eine Indexierung der in der vorangegangenen Fünfjahresperiode im SMJ erschienenen Artikel enthält. In Summe wurden so 38 Ausgaben mit insgesamt 216 wissenschaftlichen Artikeln in die Literaturreview einbezogen.

2. Das Academy of Management Journal (AMJ): Das Academy of Management Journal erscheint zweimonatlich und umfasst pro Ausgabe ab dem Jahrgang 51 neun bis 11 wissenschaftliche Artikel. Das AMJ wird seit 1957 aufgelegt. Die herausgebende Academy of Management beschreibt die Mission des AMJ folgendermaßen:

⁸² Vergleiche bspw. Schulze, Warning & Wiermann (2008) für eine Übersicht über verschiedene Rankings betriebswirtschaftlicher Zeitschriften, oder das Ranking des Vereins für Socialpolitik (www.forschungsmonitoring.org). Die Fokussierung auf englischsprachige Artikel ist keiner Gering-schätzung der deutschen Forschungslandschaft geschuldet, es gibt, entsprechend den oben genannten Ratings keine deutschsprachigen Journals im Bereich „A“ oder „A+“. Zur Rolle deutscher Wissenschaftler in der Strategischen Managementforschung vergleiche auch die Analyse zur Herkunft von Autoren in Anhang A.3.1.

⁸³ Vergleiche <http://strategicmanagement.net/pub/smj.php>.

⁸⁴ Das tatsächliche Erscheinungsdatum des SMJ läuft dem nominellen einige Monate nach, so erschien die letzte betrachtete Ausgabe, August 2009, Ende Juni 2009.

*The mission of the Academy of Management Journal is to publish empirical research that tests, extends, or builds management theory and contributes to management practice. All empirical methods -- including, but not limited to, qualitative, quantitative, field, laboratory, and combination methods -- are welcome. To be published in AMJ, a manuscript must make strong empirical and theoretical contributions and highlight the significance of those contributions to the management field. Thus, preference is given to submissions that test, extend, or build strong theoretical frameworks while empirically examining issues with high importance for management theory and practice. AMJ is not tied to any particular discipline, level of analysis, or national context.*⁸⁵

Das AMJ ist explizit auf empirische Arbeiten beschränkt. Im Gegensatz zum SMJ liegt jedoch, wie dem letzten Satz des Mission Statements zu entnehmen, kein eindeutiger Fokus auf die Strategische Managementforschung⁸⁶ vor, vielmehr wird breit von „Management“ als Erkenntnisgegenstand gesprochen, unabhängig von der Analyseebene (Person-Gruppe-Organisation-Netzwerk). Zu beachten ist, dass das AMJ durch diesen wesentlich breiteren Fokus einen deutlich geringeren Anteil an Beiträgen aufweist, die sich konkret mit der strategischen Ebene einer Organisation und Kausalitäten befassen, die einen konzeptionellen Bezug zur strategischen Ebene aufweisen. Daher ist im AMJ ein deutlich geringerer Anteil an Artikeln mit Bezug zur Leistungsoperationalisierung von Unternehmen zu erwarten.

In die Literaturreview wurden alle Ausgaben von Heft 4, Jahrgang 49 (Juli/August 2006) bis Heft 3, Jahrgang 52 (Mai/Juni 2009) einbezogen. In Summe wurden so 18 Ausgaben mit zusammen 170 Artikeln betrachtet.

Insgesamt bilden somit 386 Artikel die Datenbasis der vorliegenden Literaturreview. Der Umfang der Literaturreview bewegt sich damit im oberen Bereich vergleichbarer Studien (vergleiche Ketchen et al., 2008).

4.1.3.2 Kodierung

Die systematische qualitative Literaturreview erfordert eine konsistente Kodierung der untersuchten Artikel, auf deren Basis die der Untersuchung zu Grunde liegenden Forschungsfragen beantwortet werden können. Zur Analyse der Relevanz der Leistungsmessung wurde folgende Methodik angewandt: Zunächst las ich Überschrift und Abstract des zu untersuchenden Artikels. Stellte sich hier ein klarer Performance-Bezug heraus, erfolgte eine genaue Analyse der Methodik, des verwendeten empirischen Samples, der Discussion/Conclusion sowie der Limitations. Auf dieser Informationsbasis erfolgte dann die Einstufung in die Kategorien 1-3 zur Prüfung von Forschungsfrage 1 sowie die weitere Analyse, wie sie aus den folgenden Auswertungen erkenntlich wird. Artikel, bei denen ein Performance-Bezug in Titel und Abstract fehlte, wurden mittels Volltextsuche auf die Schlagwörter „Performance“ und „Success“ durchsucht. Treffer zogen eine genauere Textanalyse nach sich, die dann entweder zu den oben genannten Schritten führte, wenn ein empirischer Bezug zur Unternehmensperformance, etwa in Kontrollvariablen, auf diese Art und Weise sichtbar wurde, oder zur Einstufung in die Kategorie 0, wenn dies nicht der Fall war.

⁸⁵ Vergleiche <http://journals.aomonline.org/amj/home.html>.

⁸⁶ Vergleiche zur Definition und Abgrenzung des Forschungsgebietes Abschnitt 2.3.

Folgendes Flussdiagramm stellt den Ablauf schematisch dar:

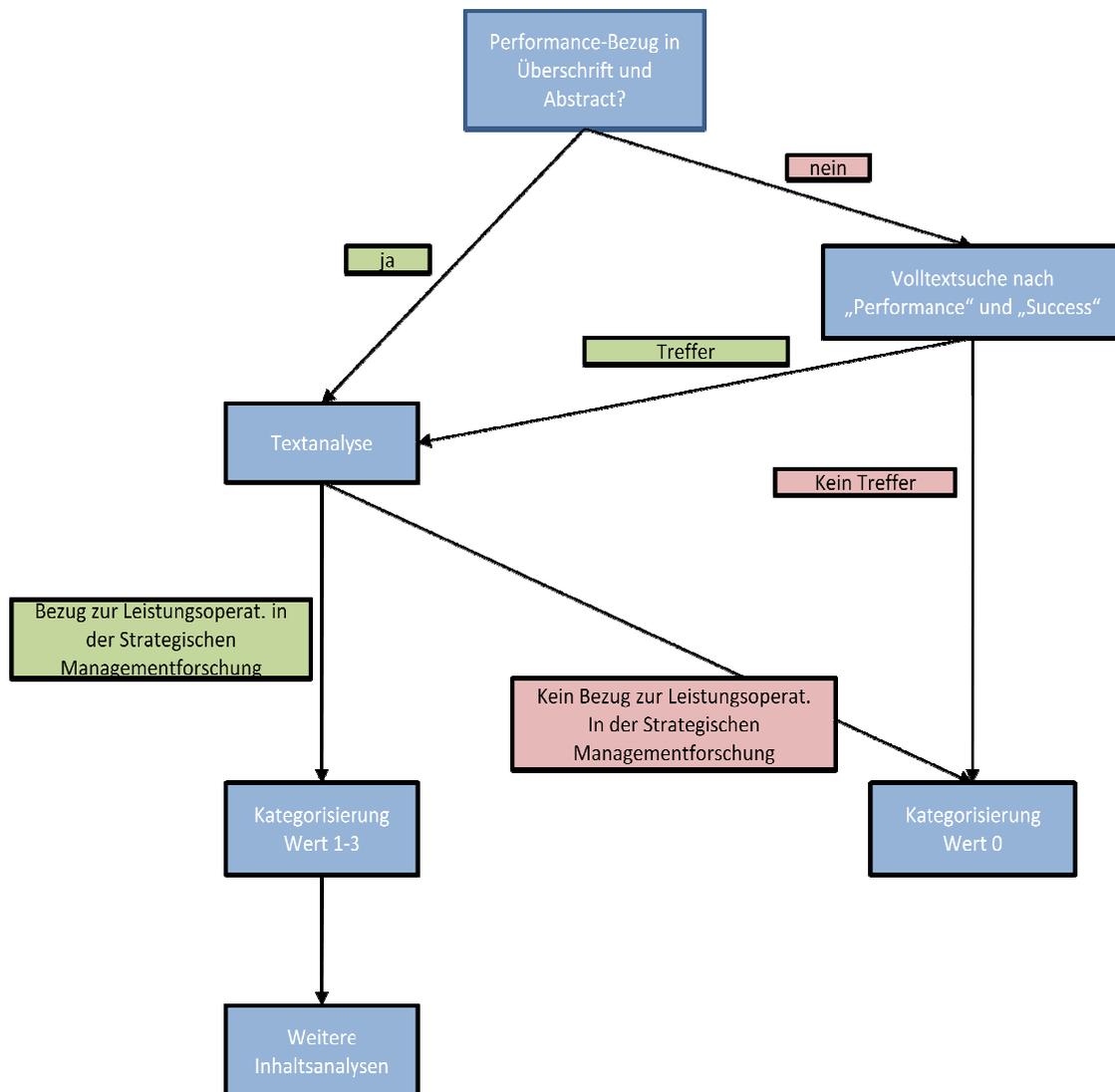


Abbildung 3: Flussdiagramm Literaturreview

Kodierung F1: Häufigkeit der Leistungsoperationalisierung in der Strategischen Managementforschung

Zur Untersuchung, wie häufig es in den im Betrachtungszeitraum erschienenen Artikeln tatsächlich notwendig ist, die Leistung einer Organisation zu operationalisieren, kodierte ich alle im Untersuchungszeitraum erschienenen Artikel auf einer nominalen⁸⁷ Skala mit Werten von 0 bis 3 hinsichtlich ihres Leistungsbezugs.

Wert 0: In diese Kategorie fallen Arbeiten, die ohne die Operationalisierung der Leistung einer Organisation auskommen, alle nichtempirischen Arbeiten sowie Labor-Settings. Ebenfalls

⁸⁷ Eine nominale Skala liegt vor, wenn die abgebildeten Werte keiner Rangfolge unterliegen, also lediglich Unterschiedlichkeit konstatiert werden kann (Hartung et al., 2009).

wurden hierunter Arbeiten gefasst, die zwar mit Leistungsoperationalisierungen arbeiten, die Methodik aber nicht dergestalt nachvollziehbar darstellen, als dass eine Bestimmung der verwendeten Operationalisierung aus dem Text des Artikels möglich wäre, etwa Wright, Kroll, Krug & Pettus (2007).

Wert 1: Diese Kategorie umfasst Beiträge, bei denen die Erfassung der Leistung nur als Kontrollvariable von Belang ist sowie Fallstudien mit Leistungsbezug. Ein Beispiel für ersteren Fall ist die Studie von Nadkarni & Barr (2008), hier untersuchen die Autoren die Wahrnehmung von Chancen durch Manager in Abhängigkeit von der Industriestruktur, befassen sich also mit der persönlichen Disposition von Managern. Die Autoren kontrollieren für unternehmensbezogene Faktoren durch die Einbeziehung der Performance-Operationalisierungen ROA, ROE und ROS.

Wert 2: Arbeiten, in denen die Operationalisierung der organisationalen Leistung bedeutsam ist, jedoch nicht die abhängige Variable darstellt. Konkret betrifft dies vier Typen von Studien: Zum Ersten solche, die einen Teilausschnitt der Leistung einer Organisation als abhängige Variable verwenden; zum Zweiten Arbeiten, die Größen operationalisieren, die erst mittelbar zu Leistung führen; zum Dritten Studien, in denen Leistung eine unabhängige Variable darstellt und schließlich Meta-Studien, deren Erkenntnisinteresse Arbeiten der Kategorie 3 umfasst. Beispiele hierfür sind Zhou et al. (2008) oder Sharfman & Fernando (2008), hier ist die Leistung der Organisation jeweils Teil einer Kausalkette, bei Sharfman & Fernando etwa werden die Kapitalkosten als abhängige Variable erfasst, die zwar eine klare Korrelation mit der Leistung einer Organisation aufweisen, zunächst jedoch nur eine Input-Variable des Unternehmens darstellen.

Wert 3: Hierunter fallen Arbeiten, in denen die Leistung der Organisation als Ganzes gemessen werden soll und in denen Leistung zudem als abhängige Variable fungiert, also Beiträge, die das erklärte oder implizite Ziel verfolgen, Leistungsunterschiede zwischen Organisationen zu untersuchen. Beispiele für Arbeiten dieser Kategorie sind etwa Ceccagnoli (2009), der Autor untersucht den Zusammenhang zwischen der Aneignungsstrategie von Innovationsergebnissen und Unternehmenserfolg, oder Mackey (2008), die den durch Heterogenität in der Person des CEOs erklärbaren Unterschied in der Leistung von Organisationen untersucht.

Nicht bewertet wurden Diskussionsbeiträge, die explizit als Replik auf einen vorab bereits publizierten Artikel oder als Klarifikation eines publizierten Artikels verfasst wurden, etwa die Diskussion im SMJ 2007 (9) zwischen Tom Connor und David Ketchen und Kollegen zum Resource- und Market-based view (Connor, 2007; Ketchen et al., 2007). Ebenfalls nicht einbezogen sind Beiträge ohne eigenen wissenschaftlichen Anspruch, etwa Vor- oder Nachworte der Herausgeber, Best-Paper-Awards und ähnliches. Im Jahrgang 50 (2007) des AMJ finden sich verschiedene Beiträge in einem Editors Forum, die sich aus einer Meta-Perspektive heraus und teilweise anekdotenhaft mit dem Fach Managementforschung beschäftigen. Aufgrund des Bruchs zur ansonsten rein empirischen Ausrichtung des Journals wurden diese Arbeiten ebenfalls nicht in die Review einbezogen.

Artikel der Kategorie 0 werden, da hier kein Bezug zur Fragestellung der Arbeit gegeben ist, in den weiteren Schritten der Literaturreview nicht mehr berücksichtigt.

Kodierung für F2: Überblick über die verwendeten Performanceoperationalisierungen:

Zur Untersuchung der verwendeten Performanceoperationalisierungen arbeite ich mit der Methode multipler Dichotomien (Brosius, 2008). Dies bedeutet, dass alle möglichen Performanceoperationalisierungen in Form einer nominalskalierten bivariaten Variablen abgebildet werden. Diese Variable nimmt den Wert 1 an, wenn die zugehörige Performanceoperationalisierung in einem Artikel verwendet wird, und den Wert 0, wenn dies nicht der Fall ist. Ich bilde also beispielsweise eine Variable „Marktkapitalisierung“, diese nimmt für alle Fälle im Datensatz, die die Leistung nicht als Marktkapitalisierung operationalisieren, den Wert 0 an, für alle Fälle, die mit Marktkapitalisierung als Leistungsoperationalisierung arbeiten, den Wert 1. Die Verwendung derartiger Dummy-Variablen bietet sich an, da sowohl Ausprägung als auch Anzahl der verwendeten Performance-Operationalisierungen unbekannt sind, die entsprechende Fragestellung also Mehrfachantworten, die Verwendung mehrerer Leistungsoperationalisierungen in einer Studie, zulässt. Die Operationalisierung mit Hilfe einer einzigen Variablen, deren Ausprägungen die möglichen Leistungsoperationalisierungen abbilden, ist daher nicht möglich, die Methode multipler Dichotomien sinnvoll.

Ich betrachte dabei parallel sowohl eine aggregierte Ebene als auch die Detailebene. Bei ersterer bilde ich vier Gruppen von Operationalisierungen: accountingbasiert, marktbasierend, prozessbasiert und expertenbasiert⁸⁸, sowie eine fünfte Kategorie „Sonstige“. Diese Systematik basiert auf einer mittels Faktorenanalyse als statistisch gehaltvoll identifizierte Kategorisierung von Leistungsindikatoren nach accountingbasiert, marktbasierend und subjektiv (in dieser Arbeit „Experten“), die von Rowe & Morrow Jr. (1999) vorgelegt wurde.⁸⁹ Ich ergänze diese Systematik um die noch fehlenden Dimensionen prozessbasiert und sonstige. Mit dieser Aufteilung bin ich in der Lage, alle vorkommenden Performance-Operationalisierungen zuzuordnen. Accountingbasierte Kennzahlen sind solche, die auf Größen der Gewinn- und Verlustrechnung aufbauen. Operationalisierungen mit auf Börsenwerten abstellendem Designatenteile ich grundsätzlich der Kategorie marktbasierend zu. Hierunter fallen insbesondere auch Größen wie Tobins q oder Jensens Alpha, die marktbasierende Designata mit accountingbasiererten mischen. Der Grund hierfür ist die Wirkung auf die Sample-Selektion, die mit der Verwendung marktbasierter Kennzahlen einhergeht. Immer wenn ein marktbasierendes Designatum Verwendung findet, ist die entsprechende Operationalisierung nur noch auf börsennotierte Unternehmen anwendbar und die Leistungsoperationalisierung damit als marktbasierend einzustufen. Die Kategorie prozessbasiert umfasst alle Operationalisierungen, die sich mit dem unmittelbaren Leistungserstellungsprozess befassen, etwa Umsatz, Mitarbeiterzahl und so weiter, sowie solche, die unabhängig von der Einschätzung von Experten vorliegen und nicht auf Größen der Gewinn- und Verlustrechnung oder börsenwertbezogenen Indikatoren basieren, etwa die Produktivität der Mitarbeiter oder die Verweildauer von Mitarbeitern im Unternehmen. Expertenbasierte Operationalisierungen arbeiten mit einer Befragung von Key Informants zu leistungsbezogenen Themen. In der Regel werden Experten mittels Fragebogen zu einer groben absoluten oder relativen Kategorisierung der Leistung ihrer Organisa-

⁸⁸ Die Frage, ob es sich bei „Experten“ tatsächlich um eine eigenständige Performanceoperationalisierung oder nicht lediglich einen anderen Weg des Zugangs bzw. der Datenakquise handelt, wird im weiteren Gang der Arbeit noch ausführlicher diskutiert.

⁸⁹ Für einen Überblick über andere gebräuchliche Gruppierungen von Leistungsoperationalisierungen vergleiche auch Combs et al. (2005).

tion aufgefördert. Sonstige umfasst Leistungsoperationalisierungen, die nur in einer spezifischen Branche einsetzbar sind. Hierunter fällt etwa der Anteil gewonnener Spiele bei Analysen im Profi-Sport (Sirmon, Gove & Hitt, 2008 oder Barden & Mitchell, 2007) oder der Anteil der Kunden einer Bank, die ihre Geschäfte online abwickeln (Weigelt & Sarkar, 2009). Ebenfalls in diese Kategorie fallen Meta-Analysen und Fallstudien. Der vorliegende Datensatz umfasst 386 Beobachtungen, darunter drei Meta-Analysen und vier Fallstudien.

Im Rahmen der Untersuchung der disaggregierten Ebene ist es sinnvoll, stark verwandte Operationalisierungen zusammenzufassen. Dies ist notwendig, um die Verteilung der Häufigkeit realistisch einschätzen zu können: gleiche „Messideen“ sollen als eine Art der Leistungsmessung dargestellt werden.

Es wurden folgende Zusammenfassungen vorgenommen:

- Die Kapitalrentabilitäten ROE, ROI, ROCE und ROA wurden aufgrund des gleichen dahinter stehenden Gedankens zu einer Variablen zusammengefasst. Alle vier messen buchhalterischen Gewinn im Verhältnis zu eingesetztem Kapital.⁹⁰ Die Umsatzrentabilität (ROS) dagegen wurde als eigenständige Variable belassen, da hier eine Größe der Gewinn- und Verlustrechnung, der Umsatz, als Bezugsgröße maßgeblich ist, keine Kapitalgröße.
- Die Marktkapitalisierung bei börsennotierten Unternehmen und die Marktwerteinschätzung durch VC-Geber werden zu einer Variablen zusammengefasst: Der Gesamtwert des Unternehmens in der Einschätzung von Eigenkapitalgebern.
- Die verschiedenen Varianten der Cumulative Abnormal Returns (CAR), die sich wesentlich nur durch die verwendeten Zeitfenster unterscheiden, wurden zu einer Größe zusammengefasst, da auch hier die gleiche Systematik zu Grunde liegt: Die Messung der kumulierten Veränderung des Aktienkurses in der Folge eines zu untersuchenden Ereignisses über einen eng abgegrenzten Zeitraum, meist wenige Handeltage.
- Alle innovationsbezogenen Indikatoren (Patentoutput, Einführung von Produkt- oder Prozessinnovationen und Indikatoren, die sich auf den FuE-Input beziehen) werden zu einer Variablen „Innovationsleistung“ zusammengelegt. Hagedoorn & Cloodt (2003) konnten zeigen, dass sich Indikatoren zur Messung der Innovationsleistung in hohem Maße statistisch überschneiden und entsprechend als gleichwertig betrachtet werden können.
- Tobins q, die Marktkapitalisierung in Relation zu den Wiederbeschaffungskosten aller Vermögensgegenstände, wird in nahezu allen untersuchten Studien als Market-to-Book Value angenähert, daher werden diese beiden Operationalisierungen zu einer Variablen zusammengefasst. Auch die subtraktive Verknüpfung zwischen Marktwert und Geschäftsvermögen, market value added, sowie dessen um die aktuellen Kapitalkosten ergänzte Variante economic value added werden unter dieser Variablen subsummiert.
- Unter Umsatz werden alle Größen des absatzbezogenen Wachstums, etwa auch die Eröffnung neuer Franchises (etwa bei Szulanski & Jensen, 2006), gefasst, ebenso wie

⁹⁰ Eine starke Korrelation konnte auch empirisch bestätigt werden, vergleiche Andersen, Denrell & Bettis, 2007.

relative Umsatzgrößen, etwa Umsatz in Relation zu den eingesetzten Assets in Luo (2007b).

- Ebenso werden alle Gewinngrößen (Nachsteuer-Gewinn, Earnings per Share usw.) zu einer Variablen zusammengefasst.

Grundsätzlich unterscheide ich nicht zwischen relativen und absoluten Werten (also etwa TSR und TSR in Relation zur Vergleichsgruppe), da sich diese Unterscheidung lediglich durch das Untersuchungsdesign, nicht jedoch durch grundlegend unterschiedliche Designata im Operationalisierungsprozess ergibt. Das gleiche gilt für die Dynamisierungen von Kennzahlen, also etwa Umsatz versus Umsatzwachstum. Einen vollständigen Überblick über die in dieser Review untersuchten Beiträge mit den jeweils verwendeten Performanceoperationalisierungen findet sich im Anhang in Tabelle A.3.

Kodierung für F3: Umgang mit dem Problem der Verallgemeinerbarkeit und häufig verwendete Kriterien zur Eingrenzung der Population

Auch hier umfasst der Datensatz Mehrfachantworten, die Gesamtpopulation wird mittels verschiedener, gleichzeitig angewandter Selektionskriterien auf kleinere Teilpopulationen reduziert, daher arbeite ich zur Bewertung, ob ein bestimmtes Selektionskriterium in einer Studie Verwendung findet, wieder mit bivariaten Variablen. Diese nehmen den Wert 1 an, wenn das entsprechende Selektionskriterium angewandt wurde, und den Wert 0, wenn dies nicht der Fall ist. Eine Ausnahme bildet die Zeitdimension. Da eine Begrenzung auf irgendeine Zeitspanne trivialerweise notwendig ist, nimmt die entsprechende Variable den Wert 1 an, wenn die betrachtete Zeitspanne kürzer ist als ein durchschnittlicher Konjunkturzyklus, den ich, der klassischen Arbeit von Kitchin (1923) folgend, mit drei Jahren annehme. Diese Abgrenzung ist weitgehend willkürlich und könnte auch durch „enges“ beziehungsweise „weites Zeitfenster“ ersetzt werden, begründet sich aber aus der in vielen Studien beobachteten Relevanz konjunktureller Phasen für Erklärungen von Kausalitäten in der Strategischen Managementforschung (siehe unten). Die Verwendung der Drei-Jahres-Periode bietet sich an, da auf diese Weise davon ausgegangen werden kann, dass kurzfristige Konjunkturschwankungen die Ergebnisse der Arbeit nicht mehr verzerren können.⁹¹ Zudem füge ich noch eine Variable „Land“ ein, die angibt, welche Region, beziehungsweise welche Volkswirtschaft in der Studie jeweils untersucht wurde. Die Variablen sind nominal skaliert.

Im Rahmen der Untersuchung, inwieweit Studien die Begrenztheit ihrer Aussagekraft auf die betrachtete Teilpopulation diskutieren nutze ich eine bivariat kodierte Variable. Diese nimmt den Wert 0 an, wenn die Einschränkung der Grundgesamtheit im Artikel nicht diskutiert wird, 1 wenn eine solche Diskussion erfolgt. Wenn diese Frage aus dem Kontext des betrachteten Artikels heraus nicht relevant ist wird dies als missing value gewertet. Letzteres tritt etwa bei Fallstudien auf, da hier die Begrenzung der Aussagemacht auf die untersuchten Fälle Bestandteil der Methodik ist, oder bei Meta-Analysen, die solche Relativierungen durch den Metastudiencharakter inkorporieren.

⁹¹ Die Betrachtung aktueller Konjunkturzyklen zeigt, dass diese in der Regel länger als drei Jahre dauern (vergleiche Abschnitt 5.4.2), für die Zwecke dieser Untersuchung erscheinen drei Jahre zur Unterscheidung zwischen langen und kurzen Zeiträumen dennoch brauchbar, da sie einen sinnvollen Minimalwert für einen konjunkturellen Zyklus darstellen.

Kodierung für F4: Umgang mit Zeitversatzproblem

Um die unterschiedlichen Möglichkeiten für den Umgang mit dem Zeitversatzproblem zu analysieren bilde ich drei Variablen. Die erste Variable bildet ab, inwieweit für Time-Lags kontrolliert wird und folgt hierbei dem Vorschlag von Bergh et al. (2004) für Untersuchungen zum Zeitversatzproblem. Sie ist bivariat und nimmt den Wert 1 an, wenn explizit für das Time-Lag Problem kontrolliert wird, 0, wenn dies nicht der Fall ist. Ist das Time-Lag Problem aus dem Untersuchungsdesign heraus nicht relevant oder ist keine Aussage hierzu möglich bleibt ein missing value.

Mit einer zweiten nominalen Variablen untersuche ich den Umgang mit dem Zeitproblem. Diese Variable hat sechs Ausprägungen, die die möglichen Lösungsoptionen des Time-Lag-Problems umfassen, wie sie im Rahmen von Abschnitt 3.4.2 erarbeitet wurden: Die Kodierung 1, „Zeitversatz“ wird gewählt, wenn explizit mit einem begründeten Zeitversatz gearbeitet wird. Mit der Kodierung 2, „über Leistungsoperationalisierung“, kodiere ich Arbeiten, die eine Performanceoperationalisierung zu Grunde legen, die eine begründete Erwartung für eine Kontrolle für Time-Lag Problematiken in sich trägt. Dies ist so grundsätzlich nur bei Messungen über Experten oder innovationsbezogene Größen der Fall (vergleiche Abschnitt 3.4.2). Kodierung 3, „Dummy“, wähle ich, wenn über Dummy-Variablen für das Vorliegen von Zeiteffekten kontrolliert wird. Die Kodierung 4, „Merge“, beschreibt Fälle, in denen Performanceoperationalisierungen über mehrere Jahre gemittelt oder kumuliert werden, um so das Zeitproblem in den Griff zu bekommen. Kodierung 5, „Methodik“, umfasst Fälle, die sich durch eine statistische Modellierung auszeichnen, die robust gegenüber Time-Lag Problemen ist. Dies ist bei Hazard-Modellen der Fall. Die Kodierung 0 schließlich, „unmöglich“, wähle ich, wenn ein Time-Lag Problem wahrscheinlich erscheint, die Operationalisierung der abhängigen Variablen aber so gewählt ist, dass sie die Kontrolle für Time-Lags grundsätzlich ausschließt. Dies ist insbesondere bei Ereignisanalysen der Fall, hier wird mit der unerwarteten Veränderung der Marktkapitalisierung in einem bestimmten Zeitfenster gearbeitet (CAR, vergleiche etwa Uhlenbruck, Hitt & Semadeni, 2006; Barnett & King, 2008; Kang, 2008; McNamara et al., 2008). Die Idee hinter diesen marktbasierenden Ereignisanalysen, die insbesondere in der Merger-Forschung Verwendung finden, ist, dass die Marktteilnehmer ihrerseits in der Lage sind, die zukünftigen Cash Flows, die sich aus dem Ereignis ergeben, zu erfassen und im Aktienkurs abzubilden. Die Grenzen dieses Ansatzes wurden bereits in Abschnitt 3.4.2 ausführlich diskutiert. Fälle, in denen kein Zeitversatzproblem auftritt oder keine klare Aussage möglich ist, werden als missing values geführt. Wurden in einer Studie mehrere Hypothesen getestet und eine Kontrolle für Time-Lags wurde nicht bei allen Tests durchgeführt, kodiere ich nach dem positiven Fall, wenn also bei einem von mehreren Hypothesentests einer Studie für Time-Lag Probleme kontrolliert wird, werte ich die Arbeit insgesamt als entsprechend kontrolliert. Die folgenden Ergebnisse überschätzen damit die Situation, die auftreten würde, wenn auf der Basis einzelner Hypothesentests und nicht auf der von Forschungsbeiträgen ausgewertet werden würde. Dieses Vorgehen ist hier angebracht, da die Sensibilität einzelner Autorenguppen für das Problem untersucht werden soll, nicht die Durchführung konkreter Hypothesentests.

Schlussendlich bilde ich eine dritte Variable die angibt, inwieweit eine Kontrolle für Time-Lags mit den entsprechenden Datensätzen grundsätzlich möglich gewesen wäre. Diese nimmt den Wert 0 an, wenn eine entsprechende Kontrolle nicht möglich ist, 1, wenn Längsschnittdaten-

sätze über längere Zeiträume vorliegen, die eine Kontrolle über gelagte Beziehungen zwischen der Variablen oder über Dummies erlauben würden, 2, wenn die Performanceoperationalisierung eine Prognosefähigkeit erwarten lässt (siehe oben). Auch hier ergeben sich missing values, wenn die Frage irrelevant für die Problemstellung oder keine Aussage möglich ist.

4.1.4 Limitationen

Eine Schwäche der hier verwendeten Methodik liegt in der subjektiven Bewertung der einbezogenen Beiträge. In der wissenschaftlichen Praxis wird bei derartigem Methodenaufbau gewöhnlich ein Vier-Augen-Prinzip angewandt, um entsprechende Abweichungen einzugrenzen. Der Grad der Abweichungen zwischen den Einschätzungen der beteiligten Wissenschaftler, die „intercoder reliability“, kann als Qualitätsindikator verstanden werden (vergleiche beispielsweise Combs et al., 2006, Orwin, 1994, oder Zott & Amit, 2008). Aufgrund des spezifischen Charakters dieser Dissertationsschrift, die als Einzelarbeit anzufertigen ist, war dies nicht möglich. Dies stellt die zentrale Limitation des vorliegenden Untersuchungsdesigns dar.

Allgemein ist die Einordnung in die Ausprägungen der verschiedenen verwendeten nominalskalierten Variablen nicht immer eindeutig. Am Beispiel der Operationalisierung zur Forschungsfrage F1, der Häufigkeit der Leistungsoperationalisierung, soll dies exemplarisch dargestellt werden. Die Trennung zwischen den Kategorien 0 bis 3 ist in der angewandten qualitativen Methodik nie frei von Interpretation. Insbesondere die Bewertung von Artikeln mit dem Wert 0 ist nicht immer einfach, da sich praktisch alle Arbeiten der Strategischen Managementforschung in irgend einer Form mit Faktoren beschäftigen, die mittelbar zur Leistung einer Organisation beitragen können. Abzugrenzen war hierbei, inwieweit eine Variable eine unmittelbare Vorstufe von „Leistung“ darstellt. Als Operationalisierung von Vorstufen einer Leistungsmessung (Wert 2) wurden Arbeiten erfasst, bei denen die Kausalität zwischen dem Beobachteten und der Leistung der Organisation anerkanntermaßen existent und positiv ist. Operationalisierungen, deren Kausalbeziehung zur Leistung einer Organisation diffus erscheint oder die Gegenstand fortgeführter Diskussionen sind, wurden mit dem Wert 0 belegt. So wurde etwa der Innovationsoutput einer Organisation als Leistungsoperationalisierung eingeordnet (Keil, Maula, Schildt & Zhara (2007) während der Internationalisierungsgrad eines Unternehmens (etwa Barkema & Shvyrkov, 2007) nicht als Leistungsindikator verstanden und der Artikel entsprechend in die Kategorie 0 eingeordnet wurde.

Diese Abschätzung ist insofern nicht unproblematisch, als dass sie zum einen nicht frei von subjektiven Wertungen ist, auch hier wäre von einer Anwendung des Vier-Augen-Prinzips eine Erhöhung der Validität der Untersuchung zu erwarten gewesen, zum anderen aber auch die Ersteller der Primärquellen unterschiedliche Einschätzungen zu Grunde legen, welche Größen als Proxy für die Performance einer Organisation wirken können. So wird die Größe einer Organisation, gemessen typischerweise am Umsatz, von vielen Autoren als Performance-Indikator genutzt (beispielsweise Somaya, Williamson & Lorinkova, 2008; Richard, Murthi & Ismail, 2007 oder Zhang, Bartol, Smith, Pfarrer & Khanin, 2008), Greve (2008) dagegen stellt die Unternehmensgröße, ebenfalls gemessen am Umsatz, als separates Ziel der Performance, gemessen anhand von accountingbasierten Indikatoren, gegenüber. In derartigen Fällen habe ich versucht, die Intention der jeweiligen Autoren nachzuvollziehen und die

Arbeiten im Geiste des Verfassers einzuordnen. Diese Unschärfe scheint vertretbar, da es im Zuge der qualitativen Literaturreview, wie in Abschnitt 4.1.2 dargelegt, lediglich um einen Überblick über das Fachgebiet, nicht aber um einen quantitativen Hypothesentest geht und zudem die Zahl der Zweifelsfälle gering ist. Kleinere Abweichungen zwischen den Kategorien sind für den Argumentationsgang der Arbeit bedeutungslos. Im Gegenzug würde eine einzel-fallbezogene Exegese des Leistungsbegriffs in den in die Review eingeflossenen Studien zu einer massiven Aufblähung der Arbeit ohne wesentlichen Erkenntniszuwachs führen.

Eine weitere Schwäche könnte in den verwendeten Suchwörtern „performance“ und „success“ liegen. Wie in Abschnitt 2.2.1 dargestellt umgibt den Leistungsbegriff auch im Englischen eine erhebliche sprachliche Unschärfe, so dass es denkbar ist, dass anstatt „performance“ beispielsweise durchgängig „profitability“ oder ähnliche Termini gebraucht worden sein könnten. Dies scheint jedoch unwahrscheinlich, geht die sprachliche Unschärfe doch eher in die Richtung, dass der Terminus „performance“ mit Bedeutungen überladen wird; dass in einem Artikel, für den die Operationalisierung von Leistung von Relevanz ist, über den gesamten Text weder das Wort „performance“ noch das Wort „success“ gebraucht wird, ist nicht zu erwarten. Stichprobenhafte Kontrollen durch genauere Textanalysen in Fällen, bei denen die Suche nach diesen Schlagworten keine Ergebnisse erbrachte, unterstützten diese Einschätzung. Nichtsdestotrotz kann davon ausgegangen werden, dass die vorliegenden Ergebnisse bezüglich Fragestellung F1 die Realität tendenziell leicht unterschätzen.

Auch die Zusammenführung eng verwandter Performanceoperationalisierungen zu umfassenderen Variablen ist im Einzelfall sicher diskutabel. So hätte etwa EVA mit einiger Berechtigung auch separat betrachtet werden können, wenngleich die in Devinney et al. (2005) zusammengetragene Evidenz zu dieser Kennzahl eher dagegen spricht. Im Einzelfall habe ich versucht, der Vorstellung der Aggregation gleicher Messideen nahe zu kommen und im Zweifelsfall eher stärker zu aggregieren. Die vorgenommenen Zusammenführungen sind dem Abschnitt 4.1.3.2 vollständig zu entnehmen, der Forderung nach interpersoneller Nachvollziehbarkeit und Wiederholbarkeit der Methodik (Green et al., 2006; Lamnek, 1995) wird damit nachgekommen.

4.2 Ergebnisse

4.2.1 Leistungsmessung in der Strategischen Managementforschung (Forschungsfrage F1)

Die folgenden Tabellen geben die Häufigkeitsverteilungen in den Kategorien 3, 2, 1 und 0 wieder. Damit wird die Frage beantwortet, wie relevant das Problem der Leistungsoperationalisierung in den aktuellen Diskursen der Strategischen Managementforschung ist.

Tabelle 2: Häufigkeiten Absolutwerte

	Kategorie 3	Kategorie 2	Kategorie 1	Kategorie 0	Summe
Beide Journals	93	63	43	187	386
Summe AMJ	28	24	15	103	170
Summe SMJ	65	39	28	84	216

Tabelle 3: Häufigkeiten Relativwerte

	Kategorie 3	Kategorie 2	Kategorie 1	Kategorie 0	Summe
Beide Journals	24%	16%	11%	48%	100%
Summe AMJ	16%	14%	9%	61%	100%
Summe SMJ	30%	18%	13%	39%	100%

Tabelle 4: Relative Häufigkeiten ohne Nullwert

	Kategorie 3	Kategorie 2	Kategorie 1
Beide Journals	47%	32%	22%
Summe AMJ	42%	36%	22%
Summe SMJ	49%	30%	21%

Die Operationalisierung von Leistung stellt, wie den obigen Tabellen zu entnehmen ist, also in der Tat ein zentrales Problem der Strategischen Managementforschung dar. In der Gesamtschau sind mehr als die Hälfte aller wissenschaftlichen Beiträge der letzten drei Jahre im untersuchten Ausschnitt der Managementforschung mit der Operationalisierung der Leistung von Organisationen befasst, richtet man das Augenmerk lediglich auf das Strategic Management Journal und damit auf den Teilausschnitt der Managementforschung, der sich explizit mit der Ebene der Organisation beschäftigt, sind es sogar 61 Prozent. Im Academy of Management Journal kehrt sich das Verhältnis zwar um, doch überrascht dies aufgrund des breiten Ansatzes dieses Journals nicht (vergleiche Abschnitt 4.1.3.1). Die Ergebnisse ändern sich nicht wesentlich nach betrachtetem Jahr.

Die Bedeutung der Leistung als Variable in der Strategischen Managementforschung wird noch unterstrichen, wenn man die Aufteilung auf die einzelnen Kategorien betrachtet: In knapp der Hälfte aller Beiträge, die sich mit der Performance von Organisationen befassen, ist diese die zentrale abhängige Variable, bei etwa einem Drittel finden Vorstufen der Leistung oder Teilleistungen von Organisationen als abhängige Variablen Verwendung, bei rund 20 Prozent der Beiträge ist die Messung von Leistung (nur) als Kontrollvariable relevant. Diese Verteilung ist zwischen den beiden betrachteten Journals weitgehend identisch. Die Leistungsoperationalisierung ist also regelmäßig ein zentraler methodischer Baustein, sehr häufig die abhängige Variable, und kann so das Ergebnis der Studien wesentlich beeinflussen.

Die Untersuchung bestätigt damit die Ergebnisse der bereits vorliegenden, ähnlich gelagerten Literaturreviews (March & Sutton, 1997; Richard, Devinney, Yip & Johnson, 2009, vergleiche Abschnitt 4.1.1). Auch wenn die Methodik aus March und Suttons Artikel nicht vollständig rekonstruierbar ist und direkte Vergleiche zu den hier vorgestellten Zahlen damit schwierig sind, lässt sich doch festhalten, dass Performance auch gut zehn Jahre nach dem Untersuchungszeitraum von March und Sutton nach wie vor eine zentrale Variable in der Strategischen Managementforschung darstellt.

4.2.2 Findet sich ein dominanter Leistungsindikator (Forschungsfrage F2)?

Betrachtet man die Bandbreite der Operationalisierungen von Leistung in der Strategischen Managementforschung, bietet sich zunächst das erwartete Bild der Methodenpluralität. Insgesamt finden 26 unterschiedliche Leistungsoperationalisierungen Verwendung.⁹² Die folgende Tabelle stellt die Häufigkeiten der aggregierten Leistungsoperationalisierungen dar.

Tabelle 5: Häufigkeit der Operationalisierungen – aggregiert

Aggregierte Leistungsoperationalisierung	Häufigkeit	Prozent	Prozent der Fälle
Accountingbasiert	63	25,9	31,7
Marktbasiert	55	22,6	27,6
Prozessbasiert	64	26,3	32,2
Expertenbasiert	37	15,2	18,6
Sonstige	24	9,9	12,1
Gesamt	243	100	122,1

N=199

In der aggregierten Betrachtung werden accounting-, markt- und prozessbasierte Größen etwa gleich häufig, in 30 Prozent der in dieser Review untersuchten Arbeiten, verwendet. Etwa jede fünfte Studie nutzt Experteneinschätzungen zur Performance-Messung, etwa jede zehnte wählt nur für ein bestimmtes Forschungssetting einsetzbare Performanceoperationalisierungen. Ebenfalls aus obiger Tabelle zu entnehmen: Im Durchschnitt arbeiteten die untersuchten Studien mit Operationalisierungen aus 1,22 Bereichen. Es zeichnet sich also in der Tat kein dominantes Paradigma der Leistungsoperationalisierung ab.

Disaggregiert man diese Daten, bietet sich das folgende Bild:

Tabelle 6: Häufigkeit der Operationalisierungen – disaggregiert

Disaggregierte Leistungsoperationalisierung	Häufigkeit	Prozent	Prozent der Fälle
ROA, ROI, ROE, ROCE	57	21,5	28,6
ROS	10	3,8	5,0
CAR	16	6,0	8,0
Market-to-Book Ratio	22	8,3	11,1
Total Shareholder Return	12	4,5	6,0
Innovationsleistung	13	4,9	6,5
Überleben/Ausscheiden	15	5,7	7,5
Umsatz	21	7,9	10,6
Interne Experten	33	12,5	16,6
Andere Operationalisierungen	24	9,1	12,1
Gesamt	265	100	133,2

N=199, nur Fälle mit n>9 aufgeführt, vollständige Tabelle im Anhang A.2

29 Prozent der untersuchten Studien arbeitet mit klassischen Kapitalrentabilitäten, jede fünfte verwendete Operationalisierung basiert auf diesen Größen. In Kombination mit der oben präsentierten relativ gleichmäßigen Verteilung auf die aggregierten Kategorien zeigt sich, dass es zwar kein dominantes Paradigma der Leistungsoperationalisierung gibt, die Bewer-

⁹² Eine vollständige Auflistung der in den untersuchten Artikeln vorzufindenden Leistungsoperationalisierungen findet sich in Tabelle A.3 im Anhang.

tung der Leistung durch das Verhältnis einer Gewinn- zu einer Kapitalgröße jedoch die bei weitem häufigste Variante im Teilbereich der accountingbasierten Operationalisierungen darstellt, also der Operationalisierungen, die auf Bilanz- und GuV-Größen abstellen. Im Rahmen der marktbasieren Kennzahlen gibt es keine solche dominante Ausprägung, bei den prozessbasierten ragen Umsatz und Überleben/Ausscheiden heraus. Im Rahmen der experientbasierten Operationalisierungen werden wesentlich mehr interne als externe Experten befragt, typischerweise das Top-Management.

Interessant ist an dieser Stelle auch, wie viele unterscheidbare Leistungsoperationalisierungen Studienergebnissen zu Grunde gelegt werden. Durch Methodenpluralität ist eine Steigerung der Aussagekraft der Ergebnisse von Studien zu erwarten, da so für etwaige Verzerrungen ungeeigneter Leistungsoperationalisierungen kontrolliert werden und ein mono-operation bias vermieden werden kann. Eine deskriptive Analyse der Anzahl der Operationalisierungen ergibt folgendes Bild:

Tabelle 7: Deskriptive Statistik Anzahl der Operationalisierungen

Zahl verwendeter Operationalisierungen	Häufigkeit	Prozent
1	145	72,9
2	42	21,1
3	12	6,0

N=199, Mittelwert=1,33

Im Durchschnitt finden also 1,33 Performanceoperationalisierungen je Studie Anwendung, in keiner Studie werden mehr als 3 Performanceoperationalisierungen verwendet. Die deutlich häufigste Ausprägung findet sich bei nur einer Operationalisierung, gut jede vierte Arbeit operationalisiert die Leistung in mindestens zwei unterscheidbaren Arten. Der Argumentation von Wortman (1994) folgend ist dieser Wert bedenklich niedrig (vergleiche auch Steigenberger, Pawlowsky, Lübcke & Schulte, 2010). Artikel im AMJ arbeiten tendenziell mit mehr Leistungsoperationalisierungen als dies in Beiträgen im SMJ der Fall ist (Durchschnitt 1,43 zu 1,28).

4.2.3 Der Umgang mit dem Problem der Verallgemeinerbarkeit (Forschungsfrage 3)

Im Rahmen von Forschungsfrage 3 sind zwei Teilfragen zu klären. Zum einen gilt es zu zeigen, nach welchen Kriterien Teilpopulationen aus der Gesamtpopulation aller Unternehmen abgegrenzt werden, zum Zweiten ist zu überprüfen, inwieweit die damit einhergehende Reduktion der Aussagekraft in qualitativ hochwertigen Arbeiten der Strategischen Managementforschung diskutiert wird. Die Betrachtung der verwendeten Selektionskriterien ergibt das folgende Bild⁹³:

⁹³ Bei den in der folgenden Tabelle fehlenden 11 Fällen handelt es sich um Fallstudien und Meta-Analysen, für die die Bestimmung von Selektionsfiltern mangels Populationsselektion nicht möglich ist. Die Ergebnisse decken sich weitgehend mit einer ähnlichen Studie, in der Bergh, Hanke, Balkundi, Brown & Chen (2004) überprüften, inwieweit für verschiedene Gefährdungen der Validität der Variablen in der Strategischen Managementforschung kontrolliert wird. Bergh und Kollegen untersuchten hierzu ein Zufallsample aus Studien, die zwischen 1994 und 1998 im Strategic Management Journal erschienen sind.

Tabelle 8: Überblick Selektionsfilter

Selektionskriterium	Häufigkeit d. Auftretens	Prozent	Prozent der Fälle
Land	174	29,8 %	92,6 %
Branche	118	20,2 %	62,8 %
Zeit	64	11,0 %	34,0 %
Notierung	77	13,2 %	41,0 %
Datenbank	59	10,1 %	31,4 %
Größe	53	9,1 %	28,2 %
Sonstige	18	3,1 %	9,6 %
Zufallsauswahl	20	3,4 %	10,6 %
Gesamt	583	100 %	310,1 %

N= 188, 11 fehlend.

Insgesamt werden in den 188 betrachteten Studien also 583 Selektionskriterien an die Gesamtpopulation angelegt, im Durchschnitt arbeitet eine Studie im Sample mit 3,1 Selektionskriterien. In fast allen Fällen werden nur bestimmte Volkswirtschaften betrachtet, in fast drei Viertel aller Fälle sind dies die USA.⁹⁴ In diesem Fall wird meist nicht reflektiert, inwieweit die Ergebnisse auf andere Länderkontexte verallgemeinerbar sind. Umgekehrt werden Abweichungen von diesem „Normalwert“ zumeist recht ausführlich begründet und analysiert (etwa bei Elbanna & Child, 2007; Mesquita & Lazzarini, 2008; Fang, Wade & Delios, 2007; David, O'Brien & Yoshikawa, 2008). Diese Fokussierung auf den westlichen Raum ist zum einen in der Dominanz US-amerikanischer Forschungsstandorte in der Managementforschung begründet (vergleiche Anhang A.3.1), wird gelegentlich aber auch durch die besonders gute Verfügbarkeit statistischer Daten in den USA erklärt (Canella Jr., Park & Lee, 2008; Shimizu, 2007).

Knapp zwei Drittel aller Studien beschränken sich auf eine bestimmte Branche, knapp die Hälfte auf börsennotierte Unternehmen. Auffällig ist, dass lediglich ein sehr geringer Teil der Arbeiten in der Strategischen Managementforschung tatsächlich eine Zufallsstichprobe aus einer bereits abgegrenzten Teilpopulation zieht. Dies ist durchaus bedenklich, wie Short et al. (2002) zeigen konnten. Die Autoren untersuchten, ähnlich wie dies im Rahmen dieser Arbeit im weiteren Verlauf noch der Fall sein wird, eine häufig beleuchtete Beziehung in der Strategischen Managementforschung, den Einfluss von CEO Duality⁹⁵ auf die Performance, und die Auswahl der Stichprobenziehung auf die ermittelten Kausalitäten und konnten hierbei zeigen, dass die Ergebnisse von Studien massiv von der Art der Stichprobenziehung abhängen. Durch eine entsprechende Sample-Selektion waren sie in der Lage, jede beliebige Richtung (positive Beziehung, keine Beziehung, negative Beziehung) zu erzeugen. Sie plädieren entsprechend für eine echte Zufallsauswahl aus der Grundgesamtheit, da nur so Verzerrungen vermieden werden könnten. In ihrer Literaturreview finden Short und Kollegen in 14 Prozent aller untersuchten Arbeiten der Strategischen Managementforschung eine Sampleselektion mit Hilfe einer Zufallsauswahl, also ähnlich geringe Werte wie sie in der hier vorliegenden Literaturreview zu Tage traten (vergleiche auch die Kommentare zu diesem Problem in Hitt et al., 2004).

⁹⁴ Eine Auswertung der geografischen Verortung von Teilpopulationen und Autoren und hieraus abgeleitete Schlussfolgerungen findet sich im Anhang A.3.1.

⁹⁵ Hiermit ist der Fall gemeint, dass CEOs auch das Board of Directors leiten, vergleiche Short, Ketchen & Palmer, 2002.

Klar ist, dass bei der verbreiteten Eingrenzung der Grundgesamtheit auf mehr oder weniger umfangreiche Teilpopulationen von einer repräsentativen Auswahl aus der Grundgesamtheit aller Unternehmen (vergleiche hierzu etwa Diekmann, 1995, 368f) keine Rede mehr sein kann. Es tritt der bereits angesprochene Zwiespalt auf: Aufgrund der Komplexität des zu untersuchenden Feldes wächst die Zahl der unerklärten Variablen massiv an, je heterogener das zu untersuchende Feld gewählt wird. Entsprechend ist zu erwarten, dass bei zahlreichen Fragestellungen aufgrund dieser unbeobachteten Effekte die Ermittlung von statistisch signifikanten Abhängigkeiten kaum mehr möglich sein dürfte. Auf der anderen Seite wird durch die Homogenisierung der Stichprobe, wie oben dargelegt, die Möglichkeit der Bewährung von Theorien erheblich reduziert. Inwieweit zwischen resultierenden Teilpopulationen Unterschiede zu erwarten sind wurde in Abschnitt 3.2.3 für die wesentlichsten Selektionskriterien diskutiert.

Zu fordern ist in jedem Fall eine ausdrückliche Würdigung, für welche Teilpopulation die gewonnenen Erkenntnisse einer Studie auf Basis der zu Grunde gelegten Daten Gültigkeit besitzen. Eine solche Explikation erfolgt relativ häufig, aber bei weitem nicht immer, wie folgende Tabelle zeigt.

Tabelle 9: Eingrenzung der Aussagemacht diskutiert?

Begrenzung der Aussagemacht	Häufigkeit	Prozent
Nicht diskutiert	92	47,9 %
Diskutiert	100	52,1 %
Gesamt	192	100 %

N = 192, fehlend: 7⁹⁶

Die Mehrheit der Studien benennt die Begrenztheit der Aussagekraft der Ergebnisse zwar, allerdings ist diese Mehrheit äußerst knapp. Es ist bedenklich, dass beinahe die Hälfte der untersuchten Artikel die erzielten Ergebnisse nicht vor dem Hintergrund der Einschränkung der Grundgesamtheit diskutieren. Die Berücksichtigung der Einschränkung der Aussagemacht sollte zudem insbesondere im Rahmen der wissenschaftlichen Weiterverwertung der Artikel im fortgeführten Diskurs Berücksichtigung finden. Dies ist regelmäßig nicht der Fall.

4.2.4 Der Umgang mit dem Time-Lag Problem

Basierend auf Forschungsfrage F4 sind zwei Untersuchungen anzustellen. Zum einen gilt es zu klären, ob die Beobachtungen von Greve & Goldeng (2004), Bergh et al. (2004) und Bergh & Holbein (1997), dass in der Strategischen Managementforschung nur sehr selten für Probleme des Zeitversatzes kontrolliert wird, immer noch aktuell sind. Alle drei Autorengruppen untersuchten dies mit Hilfe einer Literaturreview ähnlich der hier vorgelegten. Zum zweiten ist zu klären, wie in der aktuellen Forschung mit dem Problem des Zeitversatzes umgegangen wird, wenn es denn erkannt und berücksichtigt wurde.

Die folgende Tabelle gibt einen Überblick darüber, wie robust Forschungen gegenüber Zeitversatzproblemen sind.

⁹⁶ Fehlende Werte: Fallstudien und Metaanalyse, vergleiche Abschnitt 4.1.3.2.

Tabelle 10: Häufigkeit der Kontrolle für Time-Lags

Robust gegen Zeitversatzproblem?	Häufigkeit	Prozent
Ja	114	57,3 %
Nein	26	13,1 %
Keine Aussage möglich	59	29,6 %

N = 199

Entgegen der Befürchtungen von Bowen & Wiersema (1999) und Zapf et al. (1996) scheint das Zeitproblem in der wissenschaftlichen Praxis gut verstanden zu sein. Mehr als die Hälfte der untersuchten Artikel kontrollieren explizit für Time-Lags, lediglich 13 Prozent der untersuchten Artikel leiden offensichtlich unter entsprechenden Problemen.

Betrachtet man die verwendeten Methoden zur Kompensation von Zeitversatzproblemen zeigt sich, dass die in der theoretischen Diskussion für am substanzreichsten befundene Methodik, die explizite Einbeziehung eines Zeitversatzes zwischen abhängigen und unabhängigen Variablen, deutlich am häufigsten angewandt wird. Gut die Hälfte aller Studien, die überhaupt für Zeitversatz kontrollieren, arbeiten mit diesem Verfahren. Etwa ein Viertel aller Studien profitiert von Operationalisierungen der abhängigen Variablen, die ein entsprechendes Problem regelmäßig nicht erwarten lassen, die anderen Möglichkeiten des Umgangs mit Time-Lag Problemen spielen nur eine untergeordnete Rolle. Nur in 9 der 123 bewertbaren Fälle verhindert eine ungeeignete Operationalisierung der zu untersuchenden Konstrukte die Kontrolle für zu erwartende Time-Lags.

Tabelle 11: Methoden des Umgangs mit Time-Lag Problemen

Methode	Häufigkeit	Prozent
Zeitversatz	59	48,0 %
Über Leistungsoperationalisierung	32	26,0 %
Dummy-Variablen	11	8,9 %
Mittelung der abhängigen Variablen über mehrere Perioden	6	4,9 %
Hazard-Modell oder andere robuste Methode	6	4,9 %
Aus Methodik heraus unmöglich	9	7,3 %

N = 123, fehlend: 76⁹⁷

Das entsprechende Instrumentarium zur Kontrolle von Zeitversatzproblemen stellen in der Regel Panel-Datensätze zur Verfügung, wie der folgenden Tabelle zu entnehmen ist.

Tabelle 12: Optionen für Umgang mit Zeitversatzproblem

Möglichkeit zur Kontrolle für Zeitversatzeffekte	Häufigkeit	Prozent
Paneldatensatz	103	72,5 %
Prognosefähige Leistungsoperationalisierung	30	21,1 %
Aus Datenbasis heraus keine Kontrolle für Zeitversatz möglich	9	6,3 %

N = 142, fehlend: 57⁹⁸

⁹⁷ Fehlende Werte: keine eindeutige Aussage aus den in den jeweiligen Primärartikeln gegebenen Informationen möglich.

⁹⁸ Fehlende Werte: keine eindeutige Aussage aus den in den jeweiligen Primärartikeln gegebenen Informationen möglich. Die Prozentwerte addieren sich aus Gründen der Rundung nicht zu 100 % auf.

Die Verwendung von Paneldatensätzen kann damit durchaus als Standardvorgehen verstanden werden, wenn auf eine eigene Datenerhebung verzichtet wird. Den Empfehlungen von Bergh & Holbein (1997), die eben dies angemahnt hatten, wird also durchaus Folge geleistet. Die Erhebung eigener Daten erfolgt in der Regel über an Key Informants adressierte Fragebögen, ein Ansatz, dem eine gewisse Fähigkeit zur Kompensation von Zeitversatzproblemen zugeschrieben werden kann (vergleiche Abschnitt 3.4.2).

In Summe lässt sich festhalten, dass das Time-Lag Problem in dem von mir untersuchten Sample deutlich weniger dramatisch aufscheint als dies in ähnlich konzipierten Literaturreviews (etwa Bergh et al., 2004; Greve & Goldeng, 2004) der Fall war.⁹⁹ Hier ist eine Sensibilisierung der Autoren für diese Fehlerquelle und damit ein methodischer Fortschritt im Fachgebiet der Strategischen Managementforschung zu beobachten. Im Rahmen der Meta-Analyse in Kapitel 5 wird zu überprüfen sein, ob das Problem des Zeitversatzes überhaupt eine systematische Fehlerquelle für die Strategische Managementforschung darstellt. Dies wäre erwiesen, wenn sich Variationen in den Ergebnissen von Studien mit gleichem Forschungsinteresse auf unterschiedlichen Umgang mit Zeitversatzproblemen zurückführen ließen.

4.3 Diskussion

Im Rahmen der ersten der zwei dieser Arbeit zu Grunde liegenden empirischen Untersuchungen war die Frage zu klären, wie in der aktuellen Managementforschung derzeit die Leistung von Organisationen gemessen und wie mit den in Kapitel 3 entwickelten Schwierigkeiten umgegangen wird. Dem ging ich mit Hilfe einer qualitativen Literaturreview, die auf einer Vollerhebung des Academy of Management Journals und des Strategic Management Journals über jeweils drei Jahre basiert, nach: Über beide Journals arbeitet beinahe jeder zweite Artikel in irgend einer Form mit der Operationalisierung der Leistung einer Organisation, in jedem vierten der 386 untersuchten Beiträge stellt die Leistung einer Organisation die zentrale abhängige Variable dar. Diese Anteile fallen noch deutlich höher aus, wenn ausschließlich das SMJ betrachtet wird, das einen stärkeren thematischen Fokus auf die Organisationsebene aufweist. Aufbauend auf einer Kategorisierung der Performanceoperationalisierungen in accountingbasiert, marktbasierend, prozessbasiert, expertenbasiert und sonstige konnte ich zeigen, dass keine dieser fünf Hauptgruppen dominant in der Strategischen Managementforschung vertreten ist.

In Abschnitt 3.2 wurde die Notwendigkeit zur Einschränkung der Population aller Unternehmen auf eine handhabbare Teilpopulation und die damit verbundene Einschränkung der Aussagekraft von Hypothesentests diskutiert. Zugleich wurde ein Überblick über die Relevanz der wesentlichsten Kriterien, nach denen eine Eingrenzung auf Teilpopulationen in der wissenschaftlichen Praxis erfolgt, gegeben. So liegt etwa empirische Evidenz vor, dass Zusammenhänge zwischen erklärenden und erklärten Variablen in der Strategischen Management-

⁹⁹ Da die Forschungsinteressen und damit die Operationalisierungen jeweils differierten, sind die Daten aus den drei Reviews nicht unmittelbar vergleichbar. Größenordnungen lassen sich jedoch ableiten: Bergh und Kollegen konstatieren bei beinahe der Hälfte der untersuchten Beiträge Probleme in Zusammenhang mit Zeitversatz, Greve und Kollegen bemängeln, dass lediglich 19 Prozent der von ihnen untersuchten Studien mit zeitversetzten Variablen arbeiten.

forschung nicht notwendigerweise über Branchen, Unternehmensgrößenklassen und so weiter hinweg konstant sind. Hieraus leitet sich die Notwendigkeit ab, die Grenzen der Aussagekraft von Hypothesentests klar zu explizieren. Dass dies bei beinahe der Hälfte der in der Literaturreview untersuchten Studien nicht geschieht ist erstaunlich, zumal kaum angenommen werden kann, dass eine solche Relativierung der Forschungsergebnisse dem Leser auch ohne ausdrückliche Explikation ersichtlich ist, wie die Rezeption von Wissen im weiteren Diskurs der Strategischen Managementforschung zeigt. Eine wahllos herausgegriffene Zitation aus einem ebenso zufällig gewählten im Rahmen der Literaturreview betrachteten Artikel soll das Problem beispielhaft verdeutlichen: Sampson (2004) untersucht vor dem Hintergrund der Transaktionskostentheorie den Effekt von schlechter Governance auf R&D-Allianzen. Hierzu analysiert die Autorin R&D-Allianzen in der Branche Telekommunikationsausrüstung in den Jahren 1991-1993. Beachtung finden Datensätze aus zwei verschiedenen Datenbanken. Durch die Verwendung von Daten der amerikanischen Patentbehörde ergibt sich ein starker (wenn auch nicht exklusiver) Fokus auf US-Unternehmen. Sampson kann ihre Hypothese, dass sich ungeeignete Governance-Strukturen auf die Performance solcher Allianzen auswirken, bestätigen. Zitiert wird diese Studie von Aggarwal & Hsu, 2009, 845, folgendermaßen: „Sampson (2004) suggests that there are potentially important costs associated with ‘misaligned’ governance decisions that can reduce the positive benefits of collaboration.“

Dieser Satz bringt den Kern der Ergebnisse treffend zum Ausdruck, doch versäumen Aggarwal und Hsu es, die begrenzte Gültigkeit des Samples, das der Studie von Sampson zu Grunde liegt, in irgendeiner Weise zu würdigen. Wie oben gezeigt, können sowohl die untersuchte Branche als auch das Zeitfenster einen signifikanten Einfluss auf die aufgedeckten Kausalitäten ausüben. Es ist keineswegs trivial anzunehmen, dass in einem anderen Zeitfenster in einer anderen Branche die gleiche Beziehung zwischen Governance-Struktur und Performance zu finden wäre. Durch die Zitation findet also eine *Entkontextualisierung* wissenschaftlicher Erkenntnisse statt.

Eine stärkere Kontextualisierung der Managementforschung wurde vor kurzem auch von Bamberger (2008) in einer Editorial Note des Academy of Management Journal gefordert. Er spricht hierbei von der Einbeziehung von Kontextfaktoren in die Managementforschung als „nothing short of a revolution“ (S. 839). Er beobachtet eine stete Zunahme von Beiträgen im AMJ, die spezifisch Kontextvariablen in Erklärungszusammenhänge einbeziehen. Notwendig wäre überdies jedoch insbesondere die erwähnte Rekontextualisierung von *Ergebnissen*¹⁰⁰, beziehungsweise die Anerkennung der Begrenzung der allgemeinen Aussagefähigkeit von Untersuchungen, die ausschließlich einen bestimmten Kontext betrachten. Eine wichtige Aufgabe im Rahmen dieser zu fordernden Rekontextualisierung kommt hierbei sowohl den Autoren wissenschaftlicher Arbeiten zu, als auch denjenigen, die auf diesen Erkenntnissen aufbauend eigene Arbeiten verfassen. Erstere müssen klar und eindeutig die ihrer Arbeit zu Grunde liegenden Gültigkeitsbereiche benennen. Dies geschieht durch eine detaillierte und nachvollziehbare Darlegung der verwendeten Teilpopulation sowie durch eine Anerkennung der Limitationen, die die Auswahl der spezifischen Teilpopulation für die Aussagemacht der Studie bedeutet. Auf der anderen Seite sind Wissenschaftler, die auf bereits vorhandenes Wissen im Sinne einer Zitation publizierter Arbeiten in ihren eigenen Studien zurückgreifen,

¹⁰⁰ Vergleiche zum Zusammenhang zwischen kontingenztheoretischen Überlegungen, auf die sich diese Argumentation im weiteren Sinne bezieht, und dem Fachgebiet der Strategischen Managementforschung auch Hitt, Boyd & Li, 2004.

zu einem sorgfältigen Umgang mit den zitierten Quellen angehalten, insbesondere zu einer Anerkennung eben jener Begrenzung der Erklärungsmacht der zitierten Arbeiten. Die derzeit gängige Forschungspraxis ist hier fragwürdig.

Anders dagegen der Umgang mit dem Problem des Zeitversatzes zwischen Ursache und Wirkung. Hier ist klar zu erkennen, dass die Schwierigkeit, auf die in den 1990er-Jahren verschiedene Autoren hingewiesen haben, erkannt ist und ernst genommen wird. Die allermeisten Autoren, die mit Sekundärdaten arbeiten, nutzen diese in der Form von Paneldatensätzen und kontrollieren auch für etwaige Zeitversatzprobleme. Hier bestehen wenig Bedenken gegen die derzeitige Forschungspraxis.

Im Folgenden gilt es nun zu überprüfen, inwieweit die in Kapitel 3 entwickelten und in Kapitel 4 qualitativ beleuchteten Probleme tatsächlich für die beobachtete Heterogenität der Ergebnisse in der Strategischen Managementforschung verantwortlich zeichnen.

5. Metaanalyse

5.1 Forschungsansatz

5.1.1 Fragestellung

Im bisherigen Gang der Arbeit wurde das Problem der Heterogenität von Forschungsergebnissen in der Strategischen Managementforschung benannt, drei mögliche Erklärungen wurden herausgearbeitet:

- Die relative Verhandlungsmacht einzelner Stakeholdergruppen könnte als unbeobachtete Variable die Vergleichbarkeit von Forschungsergebnissen beeinträchtigen, falls diese Forschungen auf Leistungsoperationalisierungen aufbauen, die an unterschiedlichen Stellen im Prozess der Wertschöpfung und Wertverteilung in der Organisation ansetzen. Im Rahmen der qualitativen Literaturreview konnte ich zeigen, dass es in der Tat keinen dominanten Leistungsindikator gibt, vielmehr ist eine erhebliche Heterogenität in der Operationalisierung der abhängigen Variablen Leistung zu beobachten. Unterschiedliche Leistungsoperationalisierungen sollten, wenn die theoretische Herleitung zutrifft, systematisch unterschiedliche Zusammenhänge anzeigen.
- Da die Gesamtheit aller Unternehmen nicht in einzelnen Studien fassbar ist, sind Wissenschaftler im Fachgebiet gezwungen, mit mehr oder weniger großen Teilpopulationen zu arbeiten. Diese Teilpopulationen sind nicht notwendigerweise repräsentativ für die Grundgesamtheit aller Unternehmen. Im Rahmen der Literaturreview wurden die wesentlichen Kriterien, nach denen Teilpopulationen abgegrenzt werden, herausgearbeitet. Nun gilt es zu prüfen, ob Studien, die mit unterschiedlichen Teilpopulationen arbeiten, tatsächlich systematisch unterschiedliche Ergebnisse generieren. Aus den konzeptionellen Überlegungen in Abschnitt 3.2 heraus wäre dies zu erwarten.
- Schlussendlich wurde das Problem des unklaren zeitlichen Zusammenhangs zwischen abhängigen und unabhängigen Variablen als mögliche Erklärung für Heterogenität in den Ergebnissen im Fachgebiet diskutiert. Aus den konzeptionellen Überlegungen heraus wäre anzunehmen, dass Arbeiten, die Verfahren anwenden, die für das Zeitversatzproblem robust sind, Zusammenhänge systematisch besser aufdecken können als Studien, bei denen dies nicht der Fall ist.

Im Rahmen der folgenden Meta-Analyse ist nun empirisch zu überprüfen, inwieweit diese konzeptionellen Überlegungen tatsächlich Heterogenität in den Ergebnissen im Fachgebiet erklären können. Hierzu unterziehe ich vorliegende Forschungen zum Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen einer genaueren Analyse.

5.1.2 Metaanalyse – Methodik

Bei der Metaanalyse oder quantitativen Literaturreview handelt es sich um ein Verfahren zur Zusammenführung und weiteren Analyse von Forschungsergebnissen. Im Gegensatz zur qualitativen Literaturreview, die auf einen breiten Überblick über ein relativ heterogenes Forschungsfeld abzielt, versucht die quantitative Review mit statistischen Verfahren Rückschlüs-

se auf allgemeine Zusammenhänge in einem homogenen und eng abgegrenzten Forschungsgebiet zu ziehen. Untersuchungsobjekt der Metaanalyse sind einzelne Studien, diese werden im Rahmen der Analyse als Primärstudien bezeichnet. Diese müssen, um statistische Auswertungen zuzulassen, methodisch und thematisch homogen sein, das heißt sie müssen das gleiche Erkenntnisinteresse verfolgen und auf methodisch vergleichbare Art und Weise vorgehen. Üblicherweise kumuliert eine Metaanalyse die Erkenntnisse zu einer bestimmten Forschungsfrage, um diese so umfassender beantworten zu können (Cooper & Hedges, 1994). Möglich und in dem hier vorliegenden Ansatz notwendig ist auch die Ermittlung von Einflussfaktoren auf diese Ergebnisse, die Suche nach Gruppierungsvariablen (Hall & Tickle-Degnen, 1994).

Der erste Schritt im Rahmen der Metaanalyse ist daher Primärstudien zu identifizieren, die den interessierenden Zusammenhang abbilden. Anschließend werden diese Primärstudien kodiert, die Ergebnisse dieses Kodierungsvorgangs werden zur eigenen (Meta-) Analyse genutzt (Hedges & Vevea, 1998). Während die Kodierung im Rahmen der qualitativen Literaturreview auf Einschätzungen des Reviewers basiert, wird in der quantitativen Review das Datenmaterial der Primärstudien den weiteren Auswertungen zu Grunde gelegt. Auf diese Weise entsteht eine statistisch belastbare Synthese des Forschungsstandes zu einer genau definierten Forschungsfrage. Die methodischen Möglichkeiten im Rahmen der Kodierung und statistischen Weiternutzung vorliegender Forschungsergebnisse reichen hierbei von einfachen Zählmodellen (Beziehung bestätigt vs. Beziehung nicht bestätigt) über gewichtete Vergleiche bis zu Verfahren, die die Stärke von Effekten zu bestimmen in der Lage sind (Cooper & Hedges, 1994). Zu Möglichkeiten und Grenzen der Metaanalyse vergleiche auch Hall & Tickle-Degnen (1994).

Während im Rahmen der qualitativen Literaturreview die Ab- und Ausgrenzung von Artikeln aufgrund des offenen Charakters der Methodik und der angestrebten weichen Ergebnisse zumeist keine allzu große Rolle spielt, ist die Entscheidung, welche Artikel in die Metaanalyse einbezogen werden sollen, im Rahmen des quantitativen Ansatzes zentral. Zu prüfen ist, ob eine Studie die Beziehung abbildet, die in der Metaanalyse betrachtet werden soll und die notwendigen Daten zu entnehmen sind. Nur wenn dieser Test positiv ausfällt, kann die entsprechende Primärstudie in die Metaanalyse einbezogen werden. Ein wichtiger Effekt einer Metaanalyse ist die Eliminierung von Fehlern und Schwächen der einbezogenen Primärstudien (Hunter & Schmidt, 2007; Hartung, Elpelt & Klöser, 2009). Verschiedene mathematische Verfahren wurden hierzu entwickelt, womit die Methodik der Metaanalyse als gegen methodische Fehler relativ robust gilt. Die beiden wichtigsten dieser Korrekturverfahren werden im Folgenden in Abschnitt 5.4.1 einbezogen, weitere in Anhang A.3.2.1 diskutiert.

Das Sample einer Metaanalyse muss nicht notwendigerweise vollständig sein, also alle vorliegenden Primärquellen zur interessierenden Beziehung umfassen. Allerdings sollte es, wenn Vollständigkeit nicht erreicht wird, repräsentativ sein, das heißt die Auswahl des Samples muss das Ergebnis eines Zufallsprozesses oder Quota-Verfahrens sein und die wesentlichen Ausprägungen der einbezogenen Variablen umfassen (Hunter & Schmidt, 2007; Lamnek, 1995).¹⁰¹ Allerdings ist die Frage der Stichprobenziehung aus der Gesamtheit der Primärstudi-

¹⁰¹ Ist dies nicht der Fall ist eine Metaanalyse zwar immer noch möglich, allerdings sind dann andere statistische Verfahren, random effects Modelle, notwendig (Hedges & Vevea, 1998; Hunter & Schmidt, 2007).

en zumeist rein akademisch, die Anwendung statistischer Verfahren setzt gewisse Fallzahlen voraus, diese zu erreichen erfordert in der Regel eine möglichst vollständige Erschließung des interessierenden Fachgebietes unter Einbeziehung aller verfügbarer Studien, die den methodischen Mindeststandards entsprechen.

Die Ergebnisse der einbezogenen Primärstudien bilden sodann die abhängige Variable einer Meta-Analyse. Im einfachsten Fall ist diese nominal bivariat skaliert (etwa bei Newbert, 2007), man zählt also, wie viele Primärstudien den interessierenden Zusammenhang bestätigt und wie viele ihn nicht bestätigt haben und zieht hieraus Schlüsse (vergleiche zur Methodik Bushman, 1994). Newbert untersucht in seiner Metaanalyse beispielsweise unter anderem den Zusammenhang zwischen der Leistung von Organisationen und spezifischen Ressourcen im Sinne des Resource-based view, er identifiziert 29 Primärstudien, die diesen Zusammenhang abbilden, 37 % der in diesen Primärstudien durchgeführten Tests bestätigen den Zusammenhang. Newbert folgert entsprechend, dass der Zusammenhang zwischen einzelnen Ressourcen und der Leistung von Unternehmen wohl nur schwach ausgeprägt ist. Diese Zählverfahren sind gut handhabbar, da sie wenig Voraussetzungen an die Datenbasis der Primärstudien stellen und entsprechend viele Studien einbeziehen können, allerdings ist die Möglichkeit der Nutzung statistischer Verfahren durch den nominal-bivariaten Charakter der abhängigen Variablen sehr stark eingeschränkt, der Unterschied zur qualitativen Review ist gering.

Sinnvoll ist daher, eine komplexere abhängige Variable zu bilden. In aller Regel nutzt man hierfür Pearson-Korrelationskoeffizienten (vergleiche Hunter & Schmidt, 2007). Korrelationen werden in vielen Arbeiten berichtet, werden standardisiert berechnet und hängen, anders als etwa Regressionskoeffizienten in multiplen Regressionen, nicht von der Zusammenstellung eines statistischen Modells ab. Die Intervallskalierung erlaubt unter gewissen Voraussetzungen (vergleiche Abschnitt 5.4) die Anwendung komplexerer statischer Modelle. Dieses Vorgehen stellt in der qualitativ hochwertigen Managementforschung den Regelfall dar (vergleiche etwa Chang, Rosen & Levy, 2009 oder Combs, Liu, Hall & Ketchen, 2006) und soll auch in der hier vorgelegten Arbeit Verwendung finden.

In der Forschungspraxis tritt häufig das Problem auf, dass in einer Primärstudie mehrere Korrelationen angegeben werden, wenn ein Konstrukt auf unterschiedliche Weise operationalisiert wurde oder unterschiedliche Ausprägungen desselben Konstrukts interessieren. Hier müssen sinnvolle Zusammenfassungen der Einzelkorrelationen vorgenommen werden, so dass ein resultierender Wert entsteht, der sich auf das interessierende Konstrukt als Ganzes bezieht. So untersuchen beispielsweise Guerrero & Barraud-Didier (2004) den Einfluss verschiedener HRM-Praktiken auf die Performance von Unternehmen. Bündel von HRM-Maßnahmen können in ihrer Gesamtheit als Ressource im Sinne des Resource-based view aufgefasst werden (Huselid, 1995; Combs et al., 2006). In der Metaanalyse von Crook et al. (2008) interessiert der Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Performance. Guerrero und Kollegen berichtet in ihrer Studie nun Korrelationen zwischen der Leistung der Organisation, gemessen als Gewinn, und verschiedenen HRM-Maßnahmen, die als Bündel eine Ressource im Sinne des Resource-based view darstellen. Aus der Primärstudie sind Korrelationskoeffizienten in der Höhe von 0,04 für den Zusammenhang zwischen leistungsorientierter Bezahlung und Gewinn, -0,04 für die Beziehung zwischen dem Weiterbildungsniveau und dem Gewinn zu entnehmen. Ebenso wird der Gewinn in Korrelati-

on zu umfassenden Arbeitsinhalten (-0,07), Gruppenarbeit (-0,01), flexibler Gestaltung der Arbeitszeit (-0,08), Aktienoptionsprogrammen für Mitarbeiter (-0,04) und nichtmonetären Entgeltbestandteilen (-0,08) gesetzt. Das arithmetische Mittel der Korrelationskoeffizienten beträgt -0,04, dieser Wert wird als Ausprägung der abhängigen Variablen „Ressource“ in der Untersuchung zum Zusammenhang zwischen Ressource und Performance in die Metaanalyse übernommen.¹⁰²

Im nächsten Schritt ist zu entscheiden, welche weiteren Variablen in die Metaanalyse einbezogen werden sollen. Grundsätzlich ist es nicht notwendig, überhaupt weitere Variablen einzubeziehen, die abhängige Variable selbst liefert bereits die in der Regel gewünschte Information, sie beschreibt die Stärke des Zusammenhangs, der über die Pearson-Korrelationen in den Primärstudien untersucht wurde, über alle in die Metaanalyse einbezogenen Primärstudien. Häufig, so auch im hier vorliegenden Fall, interessieren jedoch Unterschiede zwischen Gruppen von Primärstudien, hierzu ist die Einbeziehung von Gruppierungsvariablen notwendig. Die Gruppierungsvariablen teilen die einbezogenen Primärstudien in Teilgruppen, mit geeigneten statistischen Verfahren (etwa ANOVA, t-Tests) werden sodann Mittelwertvergleiche zwischen den Teilgruppen durchgeführt. Abhängige Variable dieser Mittelwertvergleiche ist jeweils der aus den Primärstudien entnommene Korrelationskoeffizient. Die relativ geringen Fallzahlen, die im Rahmen von Metaanalysen in der Regel zu beobachten sind, erhöhen die Gefahr, zufällige Zusammenhänge aufzudecken. Daher ist es ratsam, mit wenigen und deduktiv gewonnenen, das heißt theoretisch unterlegten Gruppierungsvariablen zu arbeiten (Hunter & Schmidt, 2007).

Die folgende Tabelle gibt einen Überblick über die Unterschiede zwischen quantitativen Untersuchungen in Primärstudien im Rahmen der Strategischen Managementforschung und Metaanalysen.

Tabelle 13: Überblick Methodik Metaanalyse

	Primärstudie	Metaanalyse
Erkenntnisziel	Erklärung der Varianz der abhängigen Variablen, insb. der Leistung von Unternehmen, durch erklärende Faktoren innerhalb oder außerhalb der untersuchten Unternehmen.	Zusammenführung des Kenntnisstandes zu einer bestimmten Forschungsfrage (Literaturreview). Erklärung der Varianz zwischen Studien.
Fälle / Studienobjekte	Unternehmen, ggf. auch Gruppen oder Einzelpersonen in Unternehmen oder Gruppen von Unternehmen.	Publizierte Studien.

¹⁰² Diese Poolung ist im Fall breiter Konstrukte wie der Ressourcen im Sinne des Resource-based view nahezu alternativlos, kann jedoch, wie Lipsey (1994) ausführen, auch als problematisch gesehen werden, da Informationen über Einzelbeziehungen verloren gehen. Auch konzeptionell ist die Annahme, dass heterogene Indikatoren ein gemeinsames Konstrukt „Ressource“ abbilden sollen, nicht unproblematisch, dies ist jedoch weniger ein Problem der Methodik der Metaanalyse als der Theorie des Resource-based view. Diese Schwierigkeit wird in der Literatur für gewöhnlich blumig als „apples and oranges“-Problem bezeichnet, vergleiche für eine ausführlichere Diskussion Drinkmann, 1990, 23-25 sowie Abschnitt 5.6.2 in der hier vorgelegten Arbeit.

Abhängige Variable	Typischerweise die Leistung von Unternehmen.	In Primärstudien berichtete Korrelationskoeffizienten zwischen Leistung und interessierendem erklärenden Konstrukt (bspw. Ressourcen).
Unabhängige Variablen	Vermutete Einflussfaktoren auf abhängige Variable, beispielsweise Ressourcen im Sinne des RBV.	Keine.
Gruppierungs- / Moderatorvariablen	Ggf. Kontrollvariablen wie Branche, Unternehmensgröße.	Deduktiv abgeleitete vermutete Einflüsse auf Heterogenität von Studienergebnissen.
Typische Fallzahlen	100-1.000 Beobachtungen.	30-100 Primärstudien.
Anwendbare / übliche statistische Methoden	Lineare Regression, Pfadmodelle, Hazard-Modelle, Korrelationen.	Mittelwertvergleiche zwischen Gruppen von Primärstudien (durch Gruppierungsvariablen abgegrenzt): t-Test, ANOVA, ANOVA post hoc Test.

5.2 Hypothesenentwicklung

In der vorliegenden Metaanalyse folge ich eben dieser Methodik. Ich betrachte den Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Unternehmen, ein Gebiet, das sich zum einen durch eine relativ große Zahl vorliegender Primärstudien als auch durch Heterogenität in den Ergebnissen dieser Primärstudien anbietet (vergleiche ausführlich Abschnitt 5.3). In Kapitel 3 wurden die konzeptionellen Grundlagen entwickelt, die in diesem Abschnitt nun auf ihren empirischen Gehalt getestet werden sollen. Abhängige Variable der folgenden Untersuchungen ist die Korrelation zwischen Ressourcen und Performance in den Primärstudien, die Hypothesen beziehen sich auf Gruppierungsvariablen, die diese Beziehung beeinflussen.

Das Problem der Populationsselektion

Im Rahmen der Fragestellung der Populationsselektion ist zu untersuchen, inwieweit sich bestimmte Begrenzungen der Aussagekraft, also Populationsselektionsfilter, auf die beobachteten Korrelationskoeffizienten auswirken. Wie in Abschnitt 3.2 hergeleitet ist zu erwarten, dass sich die Ergebnisse der Untersuchungen kausaler Zusammenhänge systematisch unterscheiden, je nachdem welcher Ausschnitt der Realität, also der Grundgesamtheit aller Unternehmen, betrachtet wurde. Es ist zu erwarten, dass in unterschiedlichen Teilpopulationen andere Zusammenhänge zwischen Ressourcen und Leistung vorliegen. Der Einsatz von Kriterien zur Abgrenzung von Teilpopulationen müsste entsprechend die Stärke des Zusammenhangs zwischen Ressourcen und Performance systematisch beeinflussen.

Um die Zahl der Berechnungen überschaubar zu halten, betrachte ich drei der vier häufigsten Selektionsfilter: Branche, Zeit und Notierung der in einer Primärstudie untersuchten Unternehmen (vergleiche Abschnitt 3.2.3). Aufgrund der Dominanz US-amerikanischer Daten in den Primärstudien (vergleiche Anhang A.3.1) und den damit einhergehenden sehr kleinen Gruppen für Studien, die mit Daten aus anderen Volkswirtschaften arbeiten, verzichte ich auf die Einbeziehung des Selektionsfilters Land.¹⁰³ Für den hier betrachteten Selektionsfilter Branche gilt, dass Studien, die mehrere Branchen betrachten, tendenziell eine schwächer ausgeprägte Korrelation zwischen Ressourcen und Leistung finden sollten als solche, deren Sample sich aus lediglich einer oder wenigen Branchen zusammensetzt, da Brancheneffekte als unbeobachtete Variablen störend Einfluss nehmen und damit die Korrelation zwischen Ressource und Leistung abschwächen sollten. Ebenfalls sollte sich die Varianz durch den unbeobachteten Einfluss erhöhen. Zudem ist zu erwarten, dass sich die Korrelationen zwischen Samples, die auf unterschiedlichen Branchen aufbauen, statistisch signifikant unterscheiden. Entsprechend ergeben sich die folgenden Hypothesen:

H1.1: Korrelationen zwischen Leistung und Ressourcen unterscheiden sich signifikant, je nachdem, welche Branchen betrachtet werden.

H1.2: Studien, die mehr als eine Branche umfassen, berichten geringere Korrelationen und weisen höhere Varianzen auf als Studien, die sich auf eine oder wenige Branchen konzentrieren.

H1.2 ist hierbei eine speziellere Ausformung der Hypothese H1.1. Statistische Signifikanz ergibt sich nach den angewandten Verfahren (vergleiche unten).

Ähnliche Hypothesen lassen sich auch für die Selektionsfilter Zeit und Notierung ableiten. Auch hier gilt, dass sich die berichteten Korrelationen zwischen den Teilsamples unterscheiden sollten, je nachdem, welcher Zeitabschnitt den jeweiligen Primärstudien zu Grunde lag, beziehungsweise je nachdem, ob börsennotierte oder nicht börsennotierte Organisationen untersucht wurden.

H1.3: Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, welche Zeitabschnitte in den zu Grunde liegenden Primärstudien betrachtet werden.

H1.4: Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, ob gelistete oder nicht gelistete Unternehmen untersucht werden.

Das Appropriationsproblem

Zum Appropriationsproblem ist die Hypothese klar: Die Verhandlungsmacht der Anteilseigner stellt einen unbeobachteten Einfluss dar, der den beobachtbaren Teil der Beziehung zwischen Performance und Ressourcen, die als tatsächlich gegeben angenommen wird, reduzieren sollte. Entsprechend lautet die Hypothese:

H.2 Die Korrelation zwischen Performance und Ressource fällt schwächer aus, je weiter sich die zu Grunde gelegte Performanceoperationalisierung vom Konzept der Kernperformance

¹⁰³ Vergleiche Hunter & Schmidt, 1990, 111f zur Problematik kleiner Gruppen.

*entfernt, je mehr sie also bereits Objekt des Verteilungskampfes zwischen Stakeholdergruppen war.*¹⁰⁴

Das Time-Lag Problem

Im Bereich der Time-Lag Problematik ist entsprechend den Ausführungen in Abschnitt 3.4 anzunehmen, dass der unklare zeitliche Zusammenhang zwischen erklärenden und erklärten Variablen in Primärstudien zur Verschleierung des kausalen Zusammenhangs zwischen Performance und dem Vorliegen von Ressourcen im Sinne des Resource-based view führen kann. In Abschnitt 3.4.2 wurde eine Reihe an Möglichkeiten erarbeitet, mit denen Studien das Zeitversatzproblem bewältigen können. Die Anwendung dieser Verfahren sollte dazu führen, dass höhere Korrelationskoeffizienten gemessen werden als im Falle des Einsatzes nicht robuster Verfahren. Entsprechend ergibt sich die Hypothese:

H3.1: Studien, die mit Time-Lag robusten Leistungsoperationalisierungen arbeiten, finden einen stärkeren Zusammenhang zwischen Ressourcen und Performance als Studien, deren Performanceoperationalisierungen nicht robust gegenüber Zeitversatzphänomenen sind.

Eine weitere plausible Hypothese wäre, dass der Verzicht auf die Einbeziehung von Time-Lags zu unvorhersehbaren Ergebnissen führt, da keine Informationen über Einflussfaktoren verfügbar sind, die zu dem am Messzeitpunkt beobachtbarem Leistungsergebnis geführt haben könnten. Demnach wären also unsystematisch sowohl hohe als auch niedrige Korrelationen denkbar. Durch diesen erratischen Effekt würde eine Nichteinbeziehung von Time-Lags Korrelationen zwischen anderen Variablen tendenziell zu einer Erhöhung der Varianz der Korrelationen zwischen Performance und Ressourcen führen. Entsprechend lautet eine zweite Hypothese zum Time-Lag Problem:

H3.2: Der Teil der Studien, der nicht mit für Zeitversatz robusten Verfahren arbeitet, weist eine höhere Varianz in der abhängigen Variablen auf als der Teil, der robuste Verfahren einsetzt.

Wichtig ist hier noch anzumerken, dass die Hypothesen H2, H3.1 und H3.2 voraussetzen, dass tatsächlich eine Beziehung zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen existiert, die durch die entsprechenden Moderatorvariablen (relative Verhandlungsmacht von Stakeholdergruppen beziehungsweise Zeitversatzprobleme) verschleiert werden kann. Ist dies nicht der Fall wäre zu erwarten, dass unabhängig davon, welche Teilgruppen untersucht werden, stets niedrige Korrelationen zu finden sind. Die Hypothesen wären in diesem Fall nicht abzulehnen, vielmehr wäre die betrachtete Beziehung zwischen Leistung und Ressourcen im Sinne des Resource-based view zur Beantwortung der Forschungsfragen dieser Arbeit ungeeignet. Sind in irgendwelchen Teilgruppen hohe Korrelationen zu verzeichnen, sind die Hypothesentests möglich. Bisherige Forschungen zum Resource-based view (vergleiche insbesondere Peteraf & Barney, 2003) lassen den Schluss zu, dass eine Beziehung zwischen Ressourcen und Performance gegeben sein sollte, die Hypothesentests sind also mit dem gegebenen Sample durchführbar.

¹⁰⁴ Vergleiche auch das durchgerechnete Beispiel zum Einfluss bestimmter Operationalisierungen im Rahmen der Diskussion in Abschnitt 5.6.2.

5.3 Daten und Operationalisierung

5.3.1 Datensatz

Im Rahmen der vorliegenden Metaanalyse betrachte ich Arbeiten, die den Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung einer Organisation untersuchten. Die Identifikation von Moderatorvariablen, wie sie Ziel der vorliegenden Untersuchung ist, erfordert ein möglichst umfangreiches Sample, es muss also ein Untersuchungsfeld gewählt werden, zu dem zahlreiche Primärstudien vorliegen. Dies ist im Fall des Zusammenhangs zwischen Ressourcen und Performance gegeben. Der Resource-based view ist die derzeit wohl relevanteste Theorie zur Erklärung von Performanceunterschieden (Wiggins & Ruefli, 2002; Crook et al., 2008) und ist nach den Diskussionen in den 1990er Jahren um den Tautologievorwurf und den Antworten hierauf durch Barney und Peteraf (Peteraf & Barney, 2003 und andere) sowie einigen Meta-Analysen, die seine Relevanz zur Erklärung von Performanceunterschieden zumindest in Teilen bestätigten (Crook et al., 2008; Wiggins & Ruefli, 2002; Newbert, 2007) derzeit nicht ernsthaft umstritten. Wie diese Metaanalysen zeigten, liegt eine umfangreiche Datenbasis vor, die statistisch robuste metaanalytische Untersuchungen möglich macht. Gleichzeitig zeigen die Metaanalysen zum Zusammenhang zwischen Ressourcen und Performance jedoch auch die Heterogenität in den Ergebnissen der Primärstudien, die das zentrale Erkenntnisinteresse dieser Arbeit darstellt. Untersuchungen zum Zusammenhang zwischen Ressourcen und der Leistung von Organisationen eignen sich daher gut, um die hier interessierenden Hypothesen (vergleiche Abschnitt 5.2) zu untersuchen.

Der Nachteil der Betrachtung von Primärstudien zum Resource-based view liegt in der Breite des Konstrukts Ressource. Grundsätzlich ist die Metaanalyse zur Betrachtung breiter Konstrukte durchaus geeignet (Crook et al., 2008), allerdings erhöht die Breite der Konstrukte notwendigerweise auch die Optionen für die Verfasser der Primärstudien, diese Konstrukte zu operationalisieren. Für den Wissenschaftler, der eine Metaanalyse durchführen will, bedeutet dies ein Abwägungsproblem, da sich die Vielzahl an Operationalisierungen nicht mehr immer eindeutig einem bestimmten Konstrukt zuordnen lassen (vergleiche hierzu auch Abschnitt 5.6.1). Konkret gilt es hier also zu bewerten, inwieweit ein in einer Primärstudie als erklärende Variable verwendeter Indikator tatsächlich eine Ressource im Sinne des Resource-based view misst und in der Folge, ob die entsprechende Primärstudie damit in die Metaanalyse einzubeziehen ist oder nicht.¹⁰⁵

Grundsätzlich finden in vergleichbaren Arbeiten zwei Verfahren zur Identifikation relevanter Primärstudien Verwendung: Subjektive Analyseverfahren, die eine Abschätzung der Eignung von Primärstudien für die Metaanalyse auf der Basis einer Bewertung jedes Einzelfalls anstreben, und schlagwortbasierte Suchalgorithmen, die Studien identifizieren, indem sie Titel und Abstract nach bestimmten Kombinationen von Schlagwörtern durchsuchen.

Beide Verfahren fanden im Zusammenhang mit der Erstellung von Samples für Metaanalysen zur Ressourcen-Performance-Beziehung bereits Verwendung. Crook und Kollegen (Crook et al., 2008) nutzen zur Abgrenzung des Konstrukts Ressource ein mehrstufiges subjektives Be-

¹⁰⁵ Zur Rolle von Abwägungsentscheidung als Fehlerquelle in der Metaanalyse vergleiche auch Orwin, 1994. Das in der Operationalisierung des Konstrukts Leistung liegende Problem wurde bereits ausführlich in Abschnitt 2.2 diskutiert.

wertungsverfahren. Basierend auf einer breit angelegten Schlagwortsuche wurde zunächst überprüft, ob eine in einer Arbeit gefundene Operationalisierung, die als mögliche Ressource im Sinne des Resource-based view eingeschätzt wurde, eine schwer zu imitierende Größe misst. Kriterien hierfür sind Historizität, kausale Ambiguität und soziale Komplexität. Anschließend galt es, über die Werthaltigkeit der Ressourcen zu urteilen. Diese Bewertung führten die Autoren auf der Basis einer umfassenden Betrachtung der jeweils zu Grunde liegenden Theorien und empirischen Erkenntnisse durch. Zweifelsfälle wurden von den Autoren in Diskussion gelöst. Newbert (2007) dagegen arbeitete mit einem engen schlagwortbasierten Suchalgorithmus. Er stellte ein Set aus 15 Schlagwörtern zusammen, nach denen Titel und Abstract von Beiträgen in wissenschaftlichen Fachzeitschriften durchsucht wurden. Er suchte hierbei nach Studien, die bestimmte Kombinationen von Schlagworten in Titel oder Abstract anführten. Die Treffer bezog er in seine Metaanalyse zum Resource-based view ein. Dieses Verfahren hat den Vorteil, dass es die Notwendigkeit von Abwägungsentscheidungen reduziert. Bewertungen über den Ressourcencharakter erklärender Variablen sind nicht nötig, da die restriktive Schlagwortsuche das Sample auf klar Resource-based view bezogene Artikel begrenzt; dieser Vorteil ist allerdings mit einem sehr viel kleineren resultierenden Sample erkauft, da alle Studien, die sich nicht bereits im Titel oder Abstract explizit einem Test des Resource-based view verschreiben, aussortiert werden. Im Falle der Metaanalyse von Newbert wird das Sample zu klein, um noch robuste statistische Tests durchführen zu können.¹⁰⁶

Um ein möglichst breites und qualitativ hochwertiges Sample zu generieren verzichtete ich in der im Folgenden präsentierten Untersuchung auf die Erstellung eines eigenen Datensatzes nach einem der beiden möglichen Verfahren und kombinierte stattdessen die Datensätze der beiden vorliegenden Metaanalysen von Newbert (2007) und Crook und Kollegen (2008). Das resultierende Sample analysiere ich anschließend noch einmal auf Einzelfallbasis kritisch hinsichtlich des Ressourcen-Charakters der in den Primärstudien betrachteten erklärenden Variablen und eliminiere Studien, in denen mir der Ressourcen-Charakter der betrachteten erklärenden Variablen fragwürdig erscheint. Kritisch ist hierbei insbesondere die Forderung nach Nichtimitierbarkeit einer (potentiellen) Ressource (Barney, 1991; Peteraf & Barney, 2003) sowie die im Extremfall tautologische Nähe einer Ressource zum Konstrukt Performance. So entferne ich beispielsweise die Arbeit von Balkin, Markman & Gomez-Mejia (2000) aus dem Sample von Crook und Kollegen. Balkin und Kollegen untersuchten den Einfluss der Bezahlung von CEOs auf die Innovationsneigung von Unternehmen. Die Bezahlung von CEOs ist nicht schwer imitierbar, einem jeden Unternehmen steht es frei, seine Vorstände geringer oder großzügiger, eher kurzfristig oder langfristig orientiert zu entlohnen. Hier liegt auch kein Bündel an Maßnahmen vor, dem zusammen eine schwere Imitierbarkeit ähnlich dem High Performance Worksystems-Konzept (Huselid, 1995) zugeschrieben werden könnte. Entsprechend verstehe ich die Bezahlung von CEOs nicht als Ressource im Sinne des Resource-based view und nehme den Artikel nicht in die Metaanalyse auf. An anderer Stelle lag auch ein Tautologieproblem bei in den Primärstudien getesteten Hypothesen vor, entsprechende Arbeiten übernehme ich ebenfalls nicht in den Datensatz der hier präsentierten Metaanalyse.¹⁰⁷

¹⁰⁶ Newberts Sample umfasst lediglich 32 für die Zwecke dieser Metaanalyse brauchbare Primärstudien, vergleiche Abschnitt 5.4.1.

¹⁰⁷ Details dazu, welche Studien warum aus dem Sample entfernt wurden, sind Tabelle A.2 im Anhang zu entnehmen.

Durch dieses zweistufige Vorgehen, die Übernahme aller Artikel aus beiden Metaanalysen und deren Re-Evaluierung, versuche ich, zum einen ein ausreichend umfangreiches Sample für aufwändigere statistische Verfahren zu generieren und zum anderen die Abwägungsfehler durch eine zu großzügige Aufnahme zweifelhafter Primärstudien¹⁰⁸ zu minimieren. Im Weiteren eliminiere ich noch alle Primärstudien aus dem Datensatz von Newbert, die keine Korrelationskoeffizienten zwischen Ressourcen und Performance berichten. Newbert verwendet, entgegen den Empfehlungen der Methodenliteratur (vergleiche beispielsweise Hunter & Schmidt, 2007; Drinkmann, 1990; Rosenthal, 1991) in seiner Metaanalyse ein Zählmodell, bei dem Effektstärken, wie sie meiner Untersuchung zentral zu Grunde liegen, keine Rolle spielen. Hierdurch gehen Informationen verloren, es werden jedoch auch Studien verwertbar, die keine für effektstärkenbezogene metaanalytische Verfahren notwendigen Daten, Korrelationskoeffizienten, berichten. Mein Vorgehen erfordert jedoch Pearson-Korrelationskoeffizienten, die die abhängige Variable meiner Metaanalyse darstellen. Entsprechend sind Primärstudien, die diese nicht angeben, für mich grundsätzlich nicht verwertbar und müssen aus dem Datensatz entfernt werden. Die Arbeit von Crook und Kollegen dagegen arbeitet mit Effektstärken, bezieht also nur Arbeiten mit ein, die die notwendigen Daten liefern.

Testet eine Studien verschiedene unabhängige Variablen in Zusammenhang mit Performance, die alle als Ressourcen im Sinne des Resource-based view verstanden werden können, bilde ich einen Durchschnittswert für die Korrelationen einer Studie analog den Ausführungen in Abschnitt 5.1.2. Dieses Vorgehen, das üblicherweise sowohl für unabhängige als auch für abhängige Variablen der Primärstudien Verwendung findet, beschränke ich jedoch auf die erklärenden Variablen der Primärstudien, also die Ressourcen, sowie auf konzeptionell sehr eng verwandte Leistungsoperationalisierungen. Konzeptionell klar unterscheidbare Performance-Operationalisierungen betrachte ich dagegen separat, um deren Einflüsse getrennt untersuchen zu können. Im Gegensatz zu Crook et al. (2008) ist mein Untersuchungsobjekt damit also nicht die Primärstudie, sondern die Performance-Ressourcen-Beziehung (ähnlich so auch bei Newbert, 2007). Welche Verfahren als klar unterscheidbar gelten und welche Operationalisierungen als verwandt und damit zusammenfassbar angesehen werden können ist der Diskussion in Abschnitt 4.1.3.2 zu entnehmen (vergleiche hierzu auch die Synopse in Abschnitt 5.3.3). Dieses Vorgehen wird von Rosenthal (1991) empfohlen.

Damit liegen die folgenden Daten meiner Metaanalyse zu Grunde: Aus dem Datensatz von Newbert (2007) übernehme ich 33 für die Analyse verwertbare Primärstudien. In diesen werden insgesamt 50 Beziehungen zwischen Ressourcen und Performanceoperationalisierungen getestet. Der Datensatz umfasst ursprünglich 56 Primärstudien, zwei Arbeiten sind mir nicht zugänglich, von den 54 verbleibenden Studien eliminiere ich zwei Beiträge, einer erschien wissenschaftlich wenig fundiert, der zweite weist einen sehr starken Tautologieverdacht in der Performance-Ressourcen-Beziehung auf (Korrelation zwischen Kostenvorteil und finanzieller Performance). 19 Beiträge berichten keine Korrelationen und sind damit nicht verwertbar.

Crook et al. (2008) beziehen 125 Primärstudien in ihre Betrachtungen ein. Drei Arbeiten waren mir nicht zugänglich, von den verbliebenen 122 Studien waren zwei Zweitverwertungen

¹⁰⁸ Zweifelhaft ist hier natürlich nur der Ressourcencharakter der in der Primärstudie betrachteten erklärenden Variablen, nicht die Qualität der Studie an sich.

von Datensätzen, die bereits in anderen Artikeln in die Metaanalyse einbezogen wurden, diese wurden entsprechend nicht noch einmal berücksichtigt. 30 Primärstudien wurden ebenfalls in der Metaanalyse von Newbert berücksichtigt, darunter vier Studien, die ich aus dem Sample entfernte. 26 weitere Primärstudien entfernte ich aus den oben angesprochenen methodischen Gründen. Aus dem Sample von Crook und Kollegen verbleiben somit 94 verwertbare Studien, von denen 26 auch im Sample Newberts vertreten sind.

Das resultierende kumulierte und bereinigte Sample umfasst somit exakt 100 Primärstudien und 143 Korrelationskoeffizienten zwischen Ressourcen und Performance, die, als abhängige Variable, im Folgenden Verwendung finden werden. Das Gesamtsample basiert auf 24.345 Unternehmensbeobachtungen. Dieser Umfang verortet die vorliegende Metaanalyse im oberen Drittel vergleichbarer Arbeiten.¹⁰⁹ Eine vollständige Auflistung des Datensatzes ist Literaturverzeichnis III im Anhang dieser Arbeit sowie Tabelle A.2 im Anhang zu entnehmen.

Tabelle A.4 im Anhang zeigt die Herkunft der in die Metaanalyse einbezogenen Primärstudien. Hier zeigt sich erneut die Dominanz des Strategic Management Journal sowie des Academy of Management Journal, die mit 42 beziehungsweise 37 Treffern den wesentlichen Anteil am Gesamtsample auf sich vereinen. Es folgt das Journal of Business Research mit 7 Treffern. Dies bestätigt die in Abschnitt 4.1.3 getroffene Eingrenzung der qualitativen Literaturreview auf diese beiden Journals.

5.3.2 Operationalisierung

Die abhängige Variable der Metaanalyse ist, wie diskutiert, der in den einbezogenen Primärstudien ermittelte Korrelationskoeffizient zwischen Unternehmensleistung und Ressourcen. Nun ist zu entscheiden, welche Gruppierungsvariablen in die Analyse einbezogen werden sollen. Ein wesentlicher Vorteil der Metaanalyse liegt darin, dass sie es nicht nur ermöglicht, die Ergebnisse einer Vielzahl von Einzelstudien kumuliert darzustellen, sondern auch Moderatorvariablen zu identifizieren, die den Zusammenhang zwischen den abhängigen und den unabhängigen Variablen der Primärstudien beeinflussen und so über die Ergebniszusammenführung hinausgehende Erkenntnisse zu gewinnen (Hall & Tickle-Degnen, 1994). Hierzu dienen Gruppierungsvariablen, die das Gesamtsample aufteilen, um so Unterschiede zwischen Teilen des Samples aufspüren zu können (vergleiche Abschnitt 5.1.2 sowie Hunter & Schmidt, 2007). Hierzu wird der Datensatz hinsichtlich einer Dimension, die als Moderatorvariable in Frage kommt, gruppiert, der Einfluss dieser Moderatorvariablen zeigt sich dann beim Vergleich der Gruppen. Entsprechend der zu testenden Hypothesen werden folgende Gruppierungsvariablen gebildet:

Den Selektionsfilter *Branche*¹¹⁰ kodiere ich in drei Variablen: Als String-Variable, die die einzelnen Branchen abbildet, als trivariate Variable mit den Ausprägungen technologisch anspruchsvoll, technologisch weniger anspruchsvoll und gemischt/unbekannt¹¹¹, sowie in einer

¹⁰⁹ Vergleiche zur typischen Datensatzgröße in Metastudien beispielsweise den Überblick in Boyd, Gove & Hitt (2005).

¹¹⁰ Vergleiche zum Einfluss der Branche Abschnitt 3.2.3.

¹¹¹ Verschiedene Autoren vermuten starke Einflüsse des technologischen Regimes auf Kausalanalysen zwischen verschiedenen Faktoren und der Performance von Organisationen. Die Gründe hierfür liegen etwa im angenommenenmaßen turbulenteren Marktumfeld oder der höheren Umweltdynamik, dem,

sektoralen Variablen mit den vier Ausprägungen Produktion, Handel, Dienstleistungen, sowie wiederum gemischt/unbekannt.¹¹² Diese Variablen werden alternativ verwendet, zur Größe der resultierenden Zellenbelegungen vergleiche die statistischen Auswertungen in Abschnitt 5.4, zur konzeptionellen Begründung der Kodierung Abschnitt 3.2.3.

Den Einfluss eines bestimmten *Zeitabschnittes*¹¹³ überprüfe ich mit zwei Variablen: Zum einen kodiere ich nach Jahrzehnt, in dem die Datenerhebung stattgefunden hat (1970er, 1980er, 1990er und ab 2000), zum Zweiten, um eine feinkörnigere Sortierung zu erhalten, nach Konjunkturzyklus in dem die Daten erhoben wurden. Letztere Variable bestimme ich nach der Entwicklung der Veränderung des US-amerikanischen Bruttoinlandsprodukts, als Beginn eines Zyklus definiere ich das Auftreten eines starken Wachstumsanstiegs. Es ergeben sich die folgenden Zyklen: 1973-1975, 1976-1980, 1981-1983, 1984-1987, 1988-1991, 1992-1996, 1997-2003, 2004-2009.¹¹⁴ Ist der betrachtete Zeitabschnitt länger als ein Konjunkturzyklus bzw. ein Jahrzehnt oder überschneidet er zwei Zeiträume, ordne ich ihn dem Zeitraum zu, in den der größere Teil des Erhebungszeitraums fällt. Ist der betrachtete Zeitraum länger als zwei Abschnitte, kodiere ich mit „Einordnung nicht möglich“. Wo der Zeitpunkt einer Erhebung nicht eindeutig zu entnehmen war, gehe ich von einem Zeitversatz von 5 Jahren zwischen Datenerhebung und Publikation eines Artikels aus, ein Wert, der sich aus Betrachtungen vergleichbarer Arbeiten mit Zeitpunktangabe als brauchbare Näherung anbietet.

Die *Börsennotierung* von in eine Primärstudie einbezogenen Unternehmen bilde ich mit einer trivariaten Variablen ab. Diese nimmt den Wert 1 an, wenn in einer Primärstudie ausschließlich nicht börsennotierte Organisationen untersucht wurden, 2, wenn börsennotierte im Fokus standen und 0, wenn der Datensatz gemischt oder keine Aussage aus den Angaben in der Primärstudie möglich ist. Die Häufigkeit, mit der die entsprechenden Variablenausprägungen im Sample auftreten, ist wiederum Abschnitt 5.4 zu entnehmen.

Wie in Abschnitt 3.3 entwickelt gehe ich davon aus, dass die Nähe einer Performanceoperationalisierung zum Konzept der Kernperformance einen unmittelbaren Einfluss auf die Stärke der auffindbaren Korrelationskoeffizienten zwischen Ressource und Performance in den Primärstudien ausübt. Um dies zu überprüfen bilde ich zwei Gruppierungsvariablen. Die erste umfasst die bekannte Unterteilung der Leistungsoperationalisierungen in accountingbasiert, marktbasierend, prozessbasiert und Experten, die zweite kodiert die verwendeten Performanceoperationalisierungen separat, jede auftretende Operationalisierung stellt also eine Ausprägung der Gruppierungsvariablen dar.

Entsprechend den Ausführungen in Abschnitt 3.4 gehe ich davon aus, dass einige Möglichkeiten der Operationalisierungen der abhängigen Variablen Performance grundsätzlich robust gegenüber dem Problem des Zeitversatzes zwischen erklärenden und erklärten Variablen

beziehungsweise der diese Organisationen ausgesetzt sind. Vergleiche beispielsweise Collins & Clark, 2003; Hayton, 2005 oder Zahra, Ireland & Hitt, 2000.

¹¹² Unterschiede zwischen diesen Bereichen vermuten beispielsweise Skaggs & Youndt, 2004; Peng & York, 2001 oder Chandler & Hanks, 1994.

¹¹³ Vergleiche zur Diskussion des möglichen Einflusses von Zeiträumen die Ausführungen in Abschnitt 3.2.4.

¹¹⁴ Daten des U.S. Department of Commerce, <http://bea.gov>. Aufgrund der Dominanz U.S.-lastiger Datensätze in der Strategischen Managementforschung (vergleiche Anhang A.3.1) bietet sich der Rückgriff auf U.S.-amerikanische Konjunkturdaten zur Abgrenzung kürzerer Zeitfenster an.

sind, während dies bei anderen nicht der Fall ist. Entsprechend kodiere ich Timelagresistenz als bivariate Variable.¹¹⁵ Mit 1 (Zeitversatz berücksichtigt) kodiere ich, wie in Abschnitt 3.4.2 ausführlicher diskutiert, Operationalisierungen mittels qualitativer Experteneinschätzungen¹¹⁶, über einen Zeitraum aggregierte oder gemittelte Operationalisierungen, Hazard-Modelle sowie Modellierungen, die Zeitversatz explizit berücksichtigen, also Performance und Ressource in unterschiedlichen Perioden messen. Liegt keiner dieser Fälle vor oder wurden Experten explizit nach vergangenheitsbezogenen Einschätzungen befragt, kodiere ich mit 0, „Zeitversatz nicht einbezogen“.

5.3.3 Synopse

Der folgende Abschnitt soll die Kodierung kurz an zwei Beispielen veranschaulichen. Zahra & Nielsen (2002) untersuchen die Bedeutung interner Humanressourcen für den Innovationsoutput. Die Humanressourcen werden hierbei vor dem Hintergrund des Resource-based view als Ressource im Sinne dieses Ansatzes definiert. Der Innovationsoutput wird über vier Variablen gemessen: Die Zahl neuer Produkte, die Zahl radikaler Produktneuerungen, die Dauer der Überführung einer Innovation in ein neues Produkt sowie die Zahl der Patente, die im Untersuchungszeitraum erteilt wurden. Die Daten wurden über einen Drei-Jahres-Zeitraum erhoben und gemittelt. Hagedoorn & Cloodt (2003) zeigten, dass diese Operationalisierungen des Innovationsoutputs konzeptionell und empirisch stark verwandt sind, sie können daher zu einer Größe „Innovationsoutput“ zusammengefasst werden. Die Primärstudie nutzt eine ressourcenbasierte erklärende Variable: Intern verfügbare Humanressourcen. Die in der Studie angegebene Korrelationsmatrix berichtet Korrelationen von 0,27, 0,23, 0,29 und 0,25 zwischen der erklärenden und den vier erklärten Variablen. Der geringe Unterschied zwischen den Korrelationskoeffizienten unterstützt die Ergebnisse von Hagedoorn und Cloodt, dass die Innovationsindikatoren sehr stark verwandt sind. Entsprechend mittle ich die vier Korrelationen und erhalte einen gemittelten Korrelationskoeffizienten zwischen Innovationsleistung und intern verfügbaren Humanressourcen von 0,26. Dieser Wert geht als abhängige Variable in die Metaanalyse ein. Folgende Gruppierungsvariablen werden anhand der in Abschnitt 5.3.2 dargelegten Systematik gebildet:

- Die Studie beschränkt sich nicht auf eine bestimmte Branche, entsprechend kodiere ich alle drei branchenbezogenen Gruppierungsvariablen als „gemischt/unbekannt“.
- Die Autoren messen die erklärende Variable im Jahr 1996, die erklärten über die Jahre 1997-1999. Entsprechend kodiere ich die zeitpunktbezogenen Variablen als „1990er Jahre“ beziehungsweise „Konjunkturzyklus 1997-2003“, da der größere Teil des Erhebungszeitraums, der sich von 1996 bis 1999 erstreckt, in diese Periode fällt.
- Aus der Studie sind keine Angaben darüber zu entnehmen, ob die einbezogenen Unternehmen börsennotiert sind. Die entsprechende Gruppierungsvariable enthält demgemäß die Kodierung „gemischt/unbekannt“.

¹¹⁵ Der Grund für dieses pauschalisierende Vorgehen liegt im Bestreben begründet, Abwägungs- und Bewertungsentscheidungen so weit wie möglich zu minimieren. Vergleiche zum Problem von Bewertungsentscheidungen und zur Forderung nach interpersoneller Reproduzierbarkeit von Forschungsergebnissen die Abschnitte 3.1.4, 5.2 und 5.7.1.

¹¹⁶ Inwieweit Experteneinschätzungen tatsächlich als für Zeitversatzprobleme robust gelten können wird im Folgenden weiter zu diskutieren sein, vergleiche insbesondere die Abschnitte 5.4.3.3 und 5.5.

- Die Leistungsoperationalisierung, die in der Primärstudie Verwendung findet, fällt unter die Kategorie „prozessbasiert“.
- In der zweiten Gruppierungsvariablen, die sich auf das Appropriationsproblem bezieht, kodiere ich direkt nach der Art des verwendeten Leistungsindikators „Innovationsoutput“.
- Die Mittelung der abhängigen Variablen über mehrere, auf den Erhebungszeitpunkt der unabhängigen Variablen folgende Perioden stellt ein Verfahren dar, das für Zeitversatzprobleme robust ist. Entsprechend kodiere ich die Gruppierungsvariable, die sich auf den Zeitversatz bezieht, mit „1“, „Zeitversatz berücksichtigt“.

Schlussendlich sind für den weiteren Gang der Berechnungen noch die Zahl der in die Primärstudie einbezogenen Unternehmen sowie Konstruktvaliditäten der einbezogenen Variablen wichtig. Die Studie arbeitet mit 97 Unternehmensfällen, die Konstruktvalidität (Cronbachs Alpha) der erklärenden Variablen Interne Humanressourcen beträgt 0,79. Damit sind nun der Primärstudie alle Informationen entnommen, die für die Metaanalyse notwendig sind, die Primärstudie ist als ein Fall in den Datensatz der Metaanalyse übernommen.

Die Primärstudie von O'Shaugnessy, Gedajlovic & Reinmoeller (2007) soll als zweites Beispiel dienen. Dieser Beitrag ist insofern interessant, als dass hier der Test des Resouce-based view nicht im Zentrum des Interesses steht, aber dennoch Ergebnisse hierzu präsentiert werden, die den Artikel für die Metaanalyse verwertbar machen. Damit ist er ein Beispiel für einen Artikel, der mit dem schlagwortbasierten Suchverfahren nicht aufgespürt werden kann. O'Shaugnessy und Kollegen untersucht die Konstruktion und Wahrnehmung von Corporate Social Performance (CSP), ein Sammelbecken an stakeholderorientierten Maßnahmen, die in Summe als Ressource im Sinne des Resource-based view verstanden werden können. In der Studie erfassen die Autoren neben CSP auch die Market-to-Book Ratio der untersuchten Unternehmen sowie deren Return on Assets. Diese beiden Leistungsoperationalisierungen stehen nicht im Zentrum des Interesses der Autoren, werden jedoch über Korrelationsrechnungen in einen statistischen Zusammenhang mit der Ressource CSP gebracht. Der Korrelationskoeffizient zwischen CSP und ROA beträgt 0,07, der zwischen CSP und Market-to-book value 0,14. ROA und Market-to-book value unterscheiden sich anhand der in Abschnitt 4.2.3 entwickelten Systematik deutlich, beide Korrelationskoeffizienten gehen damit separat in die Metaanalyse ein. Der Artikel von O'Shaugnessy und Kollegen liefert damit zwei Datensätze für die Metaanalyse. Die Gruppierungsvariablen:

- Die Autoren untersuchen ein über Branchen gemischtes Sample, die Branchenvariablen werden entsprechend mit „gemischt“ kodiert.
- Die Datensätze, die in der Studie Verwendung fanden, wurden in den Jahren 1997-2000 erhoben, entsprechend kodiere ich die zeitpunktbezogenen Gruppenvariablen als „1990er“ und „1997-2003“.
- Das Sample besteht lediglich aus börsennotierten Unternehmen, da nur für diese ein Market-to-book value berechenbar ist. Die Gruppierungsvariable Börsennotierung kodiere ich entsprechend mit „börsennotiert“.
- Bei ROA handelt es sich um eine accountingbasierte Kennzahl, Market-to-book ratio ist marktbasierend. Die Gruppierungsvariable für das Appropriationsproblem wird entsprechend kodiert.

- Im Rahmen der Kodierung für disaggregierte Leistungsoperationalisierungen handelt es sich im Falle der beiden Leistungsindikatoren um „Kapitalrentabilität“ beziehungsweise „Market-to-book ratio“.
- Für das Zeitversatzproblem wird in der Studie nicht kontrolliert, entsprechend ist die Studie mit „0“, „Zeitversatz nicht einbezogen“ zu kodieren.

Der Artikel baut auf einem Sample von 130 Unternehmensbeobachtungen auf, Cronbachs Alpha für die erklärende Variable CSP beträgt 0,86.

5.4 Analyse

5.4.1 Statistische Korrekturfaktoren und Vorüberlegungen

Die Methodik der Metaanalyse bieten durch ihren Überblickscharakter die Möglichkeit, statistische Fehler, die in Primärstudien auftreten, zu beseitigen und Primärstudien hinsichtlich ihrer vermuteten Qualität zu gewichten. Damit geht das Instrument der Metaanalyse über die reine Zusammenführung von Forschungsergebnissen hinaus (Hunter & Schmidt, 2007). Insbesondere drei Korrekturverfahren werden für metaanalytische Samples empfohlen, die Korrektur für Sampling Error und Error of Measurement sowie die Gewichtung der Primärstudien (Hunter & Schmidt, 2007). Diese drei zentralen Verfahren sollen im Folgenden durchgeführt, beziehungsweise diskutiert werden. Hunter und Schmidt analysieren überdies eine Reihe weiterer methodischer Standardfehler, die in Metaanalysen beseitigt werden können. Eine ausführlichere Einordnung dieser Methodenprobleme in den Kontext dieser Arbeit findet sich in Anhang A.3.2, zur Anwendung in der Strategischen Managementforschung vergleiche auch Boyd et al. (2005). Die Metaanalyse setzt im Weiteren voraus, dass die einbezogenen Primärstudien voneinander unabhängig sind, also auf unterschiedliche Samples zurückgreifen (Rosenthal, 1991). Inwieweit dies in meinem Datensatz der Fall ist wird in diesem Abschnitt ebenfalls zu diskutieren sein.

Sampling Error: Eine zentrale methodische Fehlerquelle kann der Sampling Error, die zufällige Abweichung der gezogenen Stichprobe von der Grundgesamtheit, darstellen. Die Metaanalyse bietet die Möglichkeit, diese Fehlerquelle näherungsweise zu eliminieren. Hierfür wird ein Korrekturfaktor berechnet, mit dem die abhängige Variable der Metaanalyse, in diesem Fall also der Korrelationskoeffizient zwischen Leistung und Ressource in den Primärstudien, multipliziert wird. Der Sampling Error beschreibt den nichtsystematischen Einfluss des Prozesses der Stichprobenziehung in den Primärstudien auf die Studienergebnisse. Die Stichprobe kann die Grundgesamtheit, auf die sie sich bezieht, nie exakt abbilden. Je weniger Fälle aus der Grundgesamtheit gezogen werden, desto größer ist die Wahrscheinlichkeit, dass die gezogene Stichprobe unsystematisch von der Grundgesamtheit abweicht (Hunter & Schmidt, 2007). Dies kann auch für eine Metaanalyse eine Fehlerquelle darstellen, wenn sie mit relativ wenigen Primärstudien arbeitet, die selbst eher kleine Samples verwenden. Generieren Metaanalysen sehr große Stichproben, wird der Sampling Error der Einzelstudien zunehmend irrelevanter. Bei dem in der vorliegenden Studie erreichten Fallzahlen übt der Sampling Error nur einen vernachlässigbar geringen Einfluss im Bereich von 0,001 aus und kann entsprechend ignoriert werden. Die mathematische Abschätzung des Sampling Errors findet sich in Anhang A.3.2.1.

Error of Measurement: Werden Konstrukte über Fragebögen erfasst, die mehrere Items umfassen, lassen sich Reliabilitäten berechnen, die angeben, wie gut Items tatsächlich dasselbe Konstrukt abbilden. Der Grad, indem dies nicht der Fall ist, wird als Error of Measurement bezeichnet. Hunter & Schmidt (2007) argumentieren, dass dieser Error of Measurement dazu führt, dass tatsächlich vorliegende Beziehungen nicht erfasst werden können, daher sollten Korrelationskoeffizienten um einen Korrekturfaktor modifiziert werden (vergleiche auch Anhang A.3.2.1).

In der Forschungspraxis ist die Berechnung eines Näherungswertes für diesen Korrekturfaktor notwendig, da nicht alle Arbeiten Reliabilitäten berichten¹¹⁷ (vergleiche hierzu auch Boyd et al., 2005). Entsprechend folge ich dem üblichen, von Hunter und Schmidt vorgeschlagenen Verfahren und berechne aus den wenigen angegebenen Werten einen pauschalen, auf alle betrachteten Korrelationen anzuwendenden Korrekturfaktor nach der Formel

$$r_{mc} = \frac{r_m}{\sqrt{r_{perf} \times r_{res}}},$$

wobei r_{mc} der korrigierte Wert der Korrelationskoeffizienten einer bestimmten Ressourcen-Leistungs-Beziehung ist, r_m der unkorrigierte Wert, r_{perf} der durchschnittliche Reliabilitätswert der Performanceoperationalisierungen aller Studien und r_{res} entsprechend der durchschnittliche Reliabilitätswert der Ressourcenoperationalisierungen aller Studien. Der in einer Primärstudie aufzufindende Korrelationskoeffizient r_m wird also durch einen für die gesamte Metaanalyse konstanten Faktor geteilt. Da dieser Faktor größer als 0 und kleiner 1 ist, sind die korrigierten Korrelationen größer als die unkorrigierten. Für eine ausführliche Betrachtung des allgemeinen Methodenproblems der unsystematischen Messfehler vergleiche Hunter & Schmidt, 2007, 95-112.

Insgesamt berichten 28 der 100 in die Metaanalyse einbezogenen Primärstudien Reliabilitäten der erklärenden Variablen (Ressource) und 18 Reliabilitäten der erklärten (Leistung). Die Durchschnittswerte liegen bei 0,7868 für die Ressource und 0,7967 für die Leistungsvariable.¹¹⁸ Deskriptive Statistiken zu den Korrekturfaktoren findet sich in Tabelle A.5 in Anhang A.2. Aus den Mittelwerten ergibt sich nun der Korrekturfaktor als $\sqrt{0,7868} * \sqrt{0,7967} = 0,7917$. Die Korrelationskoeffizienten müssen also durch 0,7917 geteilt werden, um den pauschal für Messfehler korrigierten Korrelationswert zu generieren. Im Weiteren arbeite ich mit den errechneten korrigierten, und damit, bezogen auf die in den Primärstudien errechneten Werte, gut 25 Prozent erhöhten, Korrelationskoeffizienten.

¹¹⁷ Zur Berechnung gebräuchlicher Reliabilitätskennzahlen vergleiche beispielsweise Brosius (2008).

¹¹⁸ Auf der Performance-Seite liegt möglicherweise eine Verzerrung vor, da fast ausschließlich Fragebogenreliabilitäten, also in der Regel die Performanceoperationalisierung Interne Experten getestet wurden. Reliabilitäten, die auf anderen Kennzahlen beruhen, sind in diesem Reliabilitätswert stark unterrepräsentiert. Dieses Problem scheint so nicht lösbar und wird in vergleichbaren Metastudien auch nicht weiter thematisiert. Der Grund hierfür ist, dass Reliabilitäten nur ermittelt werden können, wenn ein Konstrukt (etwa die Performance) mit Hilfe mehrerer unterschiedlicher Items gemessen wird. Werden quantitativ eindeutige Performanceoperationalisierungen genutzt sind dies in der Regel Einzelwerte, gegebenenfalls in der Studie verwendete alternative Leistungsoperationalisierungen (etwa die beliebte Kombination ROI und Tobins q) werden nicht als Ausprägungen des selben Konstrukts zusammengefasst (was eine Reliabilitätsrechnung möglich machen würde) sondern regelmäßig separat betrachtet. Eine Berechnung von Reliabilitäten ist also für accountingbasierte, marktbasierter und prozessbasierte Leistungsoperationalisierungen zumeist methodisch nicht möglich.

Gewichtung: Fraglich ist, inwieweit Einzelstudien gewichtet werden sollten. Während Hunter & Schmidt (2007) dies nahelegen, sehen anderen Autoren (etwa Rosenthal, 1991) von einer derartigen Empfehlung ab. Das Vorgehen in der Forschungspraxis ist uneinheitlich (vergleiche Crook et al., 2008; Combs et al., 2006; Heugens & Lander, 2009). Grundsätzlich bietet sich eine Gewichtung nach der Zahl der in die jeweilige Primärstudie einbezogenen Unternehmensbeobachtungen (n) an, da diese Zahl ein Indikator für die statistische Relevanz der in der Primärstudie errechneten und als abhängige Variable in die Metaanalyse einbezogenen Korrelationskoeffizienten gelten kann. Bei Primärstudien, die mit kleinen Fallzahlen arbeiten, ist die Gefahr größer, dass Zufallskorrelationen auftreten, die dann wiederum die Aussagekraft der Metaanalyse beschädigen können. Primärstudien mit einer Gewichtung zu versehen grenzt dieses Problem ein, indem der Ausprägung der abhängigen Variablen, die auf einer methodisch robusteren, das heißt auf einem größeren Sample basierenden Primärstudie basiert, ein größeres Gewicht zukommt. Andererseits führt eine Gewichtung auch dazu, dass einige wenige Studien das Gesamtsample dominieren können, wenn die Streuung der Anzahl der jeweils einbezogenen Unternehmensbeobachtungen groß ist. Im Kontext der vorliegenden Untersuchung spricht ebenfalls gegen eine Gewichtung, dass n stark von der Methodik abhängt. Werden Sekundärdaten verwendet ist der resultierende Datensatz einer Primärstudie üblicherweise wesentlich größer als wenn eine Studie auf Primärdaten, also eigenen Datenerhebungen, aufbaut. Eine Gewichtung mit n würde also zu einer Überrepräsentation von Primärstudien mit Sekundärdatensätzen führen. Dies ist bedenklich, da verschiedentlich festgestellt wurde, dass Sekundärdatensätze teilweise „verschmutzt“ sind, da Datenbanken, aus denen die Sekundärdaten gewonnen werden, unrichtige Einträge enthalten (Starbuck, 2004).

Betrachtet man die Sample-Zahlen der in die Metaanalyse einbezogenen Studien, fällt eine erhebliche Streuung bei der Zahl der in eine Primärstudie einbezogenen Unternehmensbeobachtungen auf, wie der folgenden Tabelle zu entnehmen ist:

Tabelle 14: Deskriptive Analyse n

	n über alle einbezogenen Primärstudien
Mittelwert	289,15
Standardabweichung	472,894
Minimum	4
Maximum	3141

Die Abbildung im Boxplot zeigt, dass in der Tat nur wenige, dafür aber sehr markante Ausreißer nach oben vorliegen, während die Mehrzahl der Artikel sich im Bereich von einigen hundert Fällen bewegt. Während die teilweise sehr niedrigen Fallzahlen und die damit einhergehende erwartete geringere statistische Wertigkeit einiger Studien (Hunter & Schmidt, 2007) eine Gewichtung geboten erscheinen lassen, sprechen die Ausreißer nach oben gegen die von Hunter & Schmidt (2007) vorgeschlagene Fallzahlengewichtung. Hier würden einige wenige Studien das Gesamtsample dominieren.

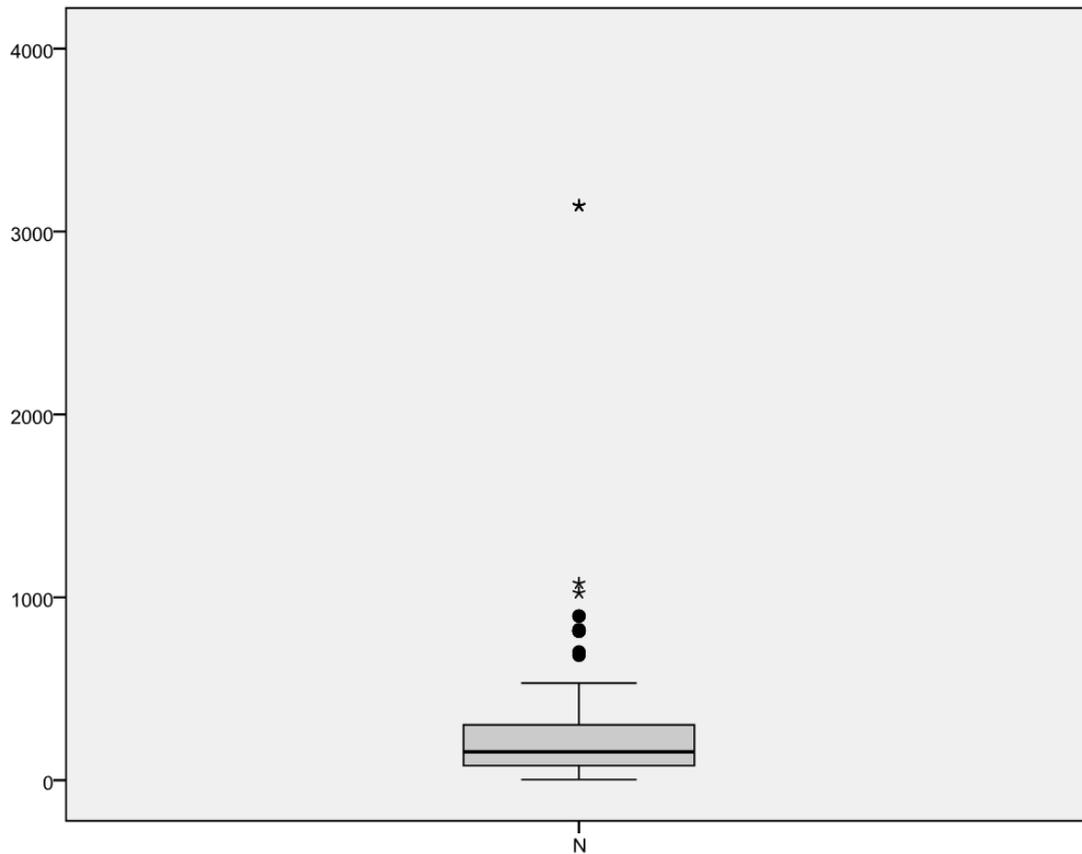


Abbildung 4: Boxplot n

Um beiden Bedenken gerecht zu werden verzichte ich auf eine Gewichtung der Studien und bereinige stattdessen den Datensatz um die fünf Fälle, die mit sehr kleinen Fallzahlen ($n < 10$) arbeiten.¹¹⁹ Das resultierende Sample umfasst so damit noch 138 Fälle und 98 Primärstudien, die ich gleichgewichtet in die weiteren Analysen einbeziehe.

Unabhängigkeit der einbezogenen Primärstudien: Da aus den in den Primärstudien verfügbaren Angaben die jeweils untersuchten Unternehmen naturgemäß nicht rekonstruierbar sind ist es nicht möglich, exakte Angaben zu möglichen Überschneidungen in den in Primärstudien gezogenen Stichproben zu machen. Eine grobe Abschätzung lässt sich anhand der Angaben zu betrachteten Branchen, Notierungsstatus der einbezogenen Unternehmen und dem Zeitraum der Stichprobenziehung jedoch erstellen. Betrachtet man die Ausprägungen der drei Gruppierungsvariablen Zeitraum (Konjunkturzyklus), Börsennotierung und Branche (String-Variable), lassen sich, nach erfolgter Bereinigung des Samples von Studien mit geringen Fallzahlen (siehe oben), 60 originär unterschiedliche Datensätze identifizieren. Dies bedeutet wohlgermerkt nicht, dass die weiteren 38 der insgesamt 98 in das Sample einbezogenen Primärstudien tatsächlich auf bereits in anderen Primäranalysen des Samples genutzten Datensätzen basieren, lediglich dass aus den vorliegenden Informationen Unterschiedlichkeit nicht eindeutig belegt werden kann. Die 60 eindeutig unterscheidbaren Primärstudien basieren auf 17.542 Unternehmensbeobachtungen. Damit ist festzuhalten, dass das Sample der Metaanalyse auf aus-

¹¹⁹ Um die Auswirkungen der kleinen Fallzahlen abzuschätzen wurden alle folgenden Berechnungen zusätzlich auch mit dem unbereinigten Sample von 143 Fällen durchgeführt, die Ergebnisse sind weitestgehend unverändert.

reichend vielen klar unterscheidbaren Unternehmensbeobachten fußt, um statistisch belastbare Untersuchungen zuzulassen.

Fortführung Synopse: In Fortführung der in Abschnitt 5.3.3 begonnenen Synopse soll die Wirkung der durchgeführten metaanalytischen Korrekturen kurz demonstriert werden.

Die Arbeit von Zahra & Nielsen (2002) ging mit einem Korrelationskoeffizienten zwischen Leistung und Ressource von 0,26 in das Sample der Metaanalyse ein. Dieser Wert wird durch den pauschalen Korrekturfaktor für Messfehler (0,7917) geteilt, der korrigierte Korrelationskoeffizient beträgt damit 0,328. Die Primärstudie basiert auf einem Sample von 97 Fällen und wird damit in das endgültige Sample übernommen.

Die Primärstudie von O'Shaughnessy et al. (2007) ging mit zwei Korrelationskoeffizienten in das Sample der Metaanalyse ein, zum einen die Korrelation zwischen Ressource (CSP) und Market-to-book value (0,14), zum zweiten die Korrelation zwischen Ressource und ROA (0,07). Dividiert durch den pauschalen Korrekturfaktor für Messfehler (0,7917) ergeben sich korrigierte Korrelationskoeffizienten von 0,177 beziehungsweise 0,088. O'Shaughnessy und Kollegen arbeiten mit 130 Unternehmensbeobachtungen, der Beitrag wird daher in das Sample der Analyse übernommen.

5.4.2 Explorative Analyse

Die folgende explorative Analyse gibt einen Überblick über den generierten Datensatz und ermöglicht zudem erste Voruntersuchungen insbesondere hinsichtlich der Verteilung der abhängigen Variablen. Zur Durchführung von ANOVA-Rechnungen ist es notwendig, dass die abhängige Variable, hier die um Messfehler korrigierten Korrelationskoeffizienten der Primärstudien, annähernd normalverteilt ist. Dies wird mit Hilfe des Kolmogorov-Smirnov-Tests überprüft.

Deskriptive Datenanalyse: Die folgende Tabelle beschreibt die abhängige Variable „korrigierte Korrelation“. Wie zu erkennen ist finden Autoren recht unterschiedliche Korrelationskoeffizienten zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen; der Mittelwert liegt bei 0,23, im allgemeinen wird also, so lassen sich diese Werte deskriptiv interpretieren, eine Beziehung zwischen Performance und Ressourcen bestätigt, eine Unterstützung der Theorie des Resource-based view.

Tabelle 15: Deskriptive Analyse der abhängigen Variablen „korrigierte Korrelation“

Mittelwert	0,2295
Standardabweichung	0,18396
Minimum	-0,20
Maximum	0,67

N = 138

Test auf Normalverteilung: Die im Weiteren durchzuführenden statistischen Verfahren (ANOVA) erfordern eine Normalverteilung der abhängigen Variablen. Ein graphischer Test auf Normalverteilung zeigt ein gutes Ergebnis. Der für größere Fallzahlen empfohlene

Kolmogorov-Smirnov-Test auf Normalverteilung¹²⁰ mit einer Signifikanzanpassung nach Liljefors prüft die Hypothese, dass die untersuchte Variable nicht normalverteilt ist. Das Signifikanzniveau von 0,2 lässt eine Annahme dieser Hypothese nicht zu. Damit kann festgestellt werden, dass die abhängige Variable mit befriedigender Näherung nicht normalverteilt ist.¹²¹ Verfahren die Normalverteilung voraussetzen können also angewandt werden.

Tabelle 16: Kolmogorov-Smirnov-Test auf Normalverteilung

Untersuchte Variable	Statistik	Freiheitsgrade	Signifikanz
Korrigierte Korrelationskoeffizienten	0,040	138	0,200

N = 138

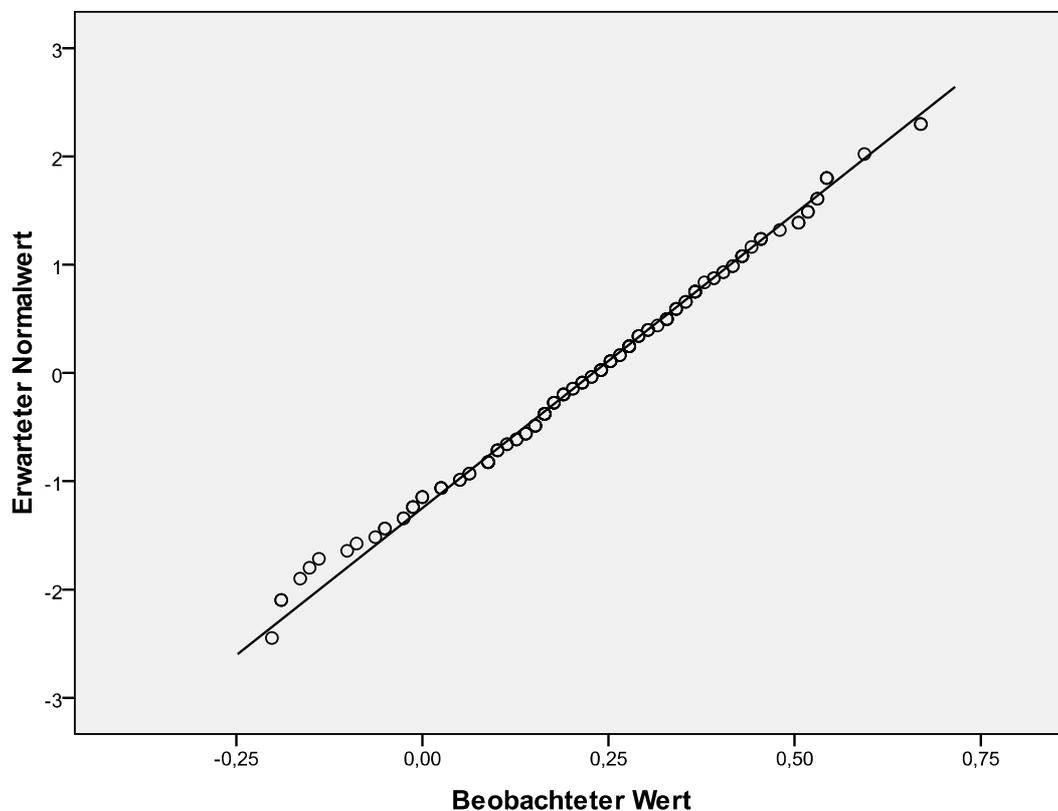


Abbildung 5: Test auf Normalverteilung der abhängigen Variablen, grafische Auswertung

5.4.3 Hypothesentests

Im folgenden Abschnitt sind nun die in Kapitel 3 deduktiv hergeleiteten und in Abschnitt 5.2 ausformulierten Hypothesen empirisch zu überprüfen. Im Folgenden werden die Ergebnisse präsentiert, eine Diskussion dieser Ergebnisse erfolgt anschließend in Abschnitt 5.5. Zunächst wird das Problem der Verallgemeinerbarkeit von Forschungsergebnissen über Teilpopulationen untersucht, hier ist die Hypothese zu testen, dass sich die Ergebnisse von Studien, die

¹²⁰ Zum Hintergrund dieses Testverfahrens vergleiche beispielsweise Hartung, Elpelt & Klösener (2009).

¹²¹ Die doppelte Verneinung ist der präzisen Wiedergabe des mathematischen Ergebnisses des Tests geschuldet.

unterschiedliche Teilpopulationen betrachten, statistisch signifikant unterscheiden. Anschließend ist zu prüfen, ob die in Primärstudien gewonnenen Korrelationskoeffizienten so von der Leistungsoperationalisierung abhängen wie von den appropriationstheoretischen Überlegungen vorhergesagt. Zum Dritten beleuchte ich schließlich, inwieweit die Nichtbeachtung des Zeitversatzproblems tatsächlich dazu führt, dass Zusammenhänge zwischen Ressourcen und Performance nicht aufgedeckt werden.

5.4.3.1 Das Problem der Verallgemeinerbarkeit über Teilpopulationen

H1.1: Korrelationen zwischen Leistung und Ressourcen unterscheiden sich signifikant, je nachdem, welche Branchen betrachtet werden.

H1.2: Studien, die mehr als eine Branche umfassen, berichten geringere Korrelationen und weisen höhere Varianzen auf als Studien, die sich auf eine oder wenige Branchen konzentrieren.

H1.3: Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, welche Zeitabschnitte in den zu Grunde liegenden Primärstudien betrachtet werden.

H1.4: Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, ob gelistete oder nicht gelistete Unternehmen untersucht werden.

Selektionsfilter: Branche (H1.1 und H1.2)

Die in die Metaanalyse einbezogenen Primärstudien untersuchen Unternehmen aus einer Vielzahl von Branchen. Die am häufigsten vertretene Einzelbranche ist die Bankbranche (5 Fälle), gefolgt von Krankenhäusern (4 Fälle). Keine weitere Branche kommt mehr als zwei Mal im Sample vor. Viele Studien beschränken sich nicht auf eine einzelne Branche sondern grenzen ihr Sample stattdessen nach Sektoren ab (produzierendes Gewerbe, Handel, etc.). Durch die resultierenden kleinen Subsamples ist ein Vergleich auf Unterschiede zwischen einzelnen Branchen nicht möglich. Möglich ist eine solche Betrachtung jedoch auf Basis des Technologieregimes, obwohl auch hier die Fallgruppen sehr klein ausfallen. Lediglich 36 Primärstudien arbeiten mit Teilpopulationen, die sich eindeutig als Lowtech-, beziehungsweise Hightech-Unternehmen klassifizieren lassen. 22 Primärstudien beschäftigen sich ausschließlich mit Lowtech-, 14 konzentrieren sich auf Hightech-Unternehmen. Die Größe der beiden Gruppen liegt damit am unteren Rand dessen, was für Gruppenvergleiche noch akzeptabel ist (Hunter & Schmidt, 2007). Um auf Unterschiedlichkeit zu testen, folge ich dem Standardverfahren der Methodenliteratur (Hunter & Schmidt, 2007; Rosenthal, 1991) und teile das Sample entsprechend der interessierenden Gruppenvariable, hier „Technologieregime“, und vergleiche die Mittelwerte der resultierenden Teilgruppen mit geeigneten statistischen Verfahren. Verfügt die Gruppierungsvariable nur über zwei Ausprägungen, wie hier der Fall, kann die Signifikanz etwaiger Unterschiede mit Hilfe eines t-Tests für unabhängige Teilstichproben überprüft werden.

Der Wert der Teststatistik, der t-Wert, berechnet sich aus Mittelwerten und Varianzen zweier Stichproben (Brosius, 2008, 468-474). Der Wert wird 0, wenn die Mittelwerte in den beiden Gruppen exakt gleich sind. Ist dies nicht der Fall, hängt der Wert von t von den Varianzen und

Samplegrößen ab. Je größer die Fallzahlen in den Teilgruppen und je kleiner die Varianzen, desto größer wird der t-Wert. Hohe Fallzahlen und geringe Varianzen deuten dabei darauf hin, dass festgestellten Unterschiede zwischen den Teilgruppenmittelwerten nicht zufällig sondern systematisch sind. Entsprechend ist bei sehr heterogenen (hohen Varianzen) oder sehr kleinen (geringen Fallzahlen) Samples die Wahrscheinlichkeit relativ hoch, dass die Abweichung der Mittelwerte der Teilstichproben lediglich Ergebnis eines nicht für die Grundgesamtheit repräsentativen Zufallsprozesses in der gezogenen Stichprobe ist, die Gruppen also nicht tatsächlich unterschiedlich sind, beziehungsweise eine Aussage über die Unterschiedlichkeit nicht mit statistischer Signifikanz getroffen werden kann. Der t-Test prüft also auf Unterschiedlichkeit von Teilstichproben, statistisch signifikante t-Werte lassen den Schluss zu, dass die Gruppen tatsächlich unterschiedlich sind (vergleiche ausführlicher Hartung et al., 2009).¹²²

Wie die folgende Tabelle zeigt, ist die Hypothese der Unterschiedlichkeit der beiden Teilgruppen Hightech und Lowtech klar abzulehnen. Die Mittelwerte der Teilstichproben unterscheiden sich lediglich marginal (0,2337 zu 0,2346), entsprechend liegt der t-Wert nahe Null. Der Signifikanzwert von 0,989 besagt, dass die Annahme der Hypothese der Unterschiedlichkeit mit knapp 99-prozentiger Wahrscheinlichkeit ein Fehler wäre.

Tabelle 17: t-Test, Gruppierungsvariable „Technologieregime“

	T-Wert	Freiheitsgrade	Signifikanzniveau	Mittlere Differenz	Standardfehler der Differenz
Korrigierte Korrelation	-0,014	30,229	0,989	-0,00090	0,06617

N = 36

Neben der Kodierung für einzelne Branchen und für technologische Regime liegt noch eine dritte branchenbezogene Variable vor, die die Branchenzugehörigkeit der in den Primärstudien verwendeten Teilpopulationen nach Sektoren kodiert – Produzierendes Gewerbe, Handel und Dienstleistungen. Da lediglich 4 Fälle Unternehmen aus dem Bereich Handel beleuchten, beschränke ich die weitere Analyse auf die beiden Gruppen Produktion und Dienstleistungen, die 47 beziehungsweise 20 mal im Sample auftreten.

Die Mittelwerte unterscheiden sich hier deutlicher (0,2201 zu 0,1914), allerdings auch nicht deutlich genug, als dass die Teststatistik signifikante Ergebnisse erbringen würde. Das Signifikanzniveau von 0,58 erlaubt es nicht, die Hypothese der Unterschiedlichkeit anzunehmen.

Tabelle 18: t-Test, Gruppierungsvariable „Branchen aggregiert“

	T-Wert	Freiheitsgrade	Signifikanzniveau	Mittlere Differenz	Standardfehler der Differenz
Korrigierte Korrelation	0,560	36,708	0,579	0,02874	0,05134

N = 67

¹²² Der entsprechende Umkehrschluss, dass ein nicht signifikanter t-Test auf Gleichheit der Teilsamples schließen lassen könnte, ist jedoch unrichtig. Niedrige t-Werte können, wie dargelegt, sowohl aus geringen Mittelwertdifferenzen als auch aus stark streuenden oder kleinen Stichproben resultieren.

Es lässt sich also festhalten, dass kein statistisch signifikanter Einfluss der Branche auf die Beziehung zwischen Leistung und Ressourcen ermittelt werden kann. Da bereits Hypothese H1.1 Ablehnung findet, ist ein Test für die weiterführende Hypothese H1.2 hinfällig.

Selektionsfilter Zeitraum

Es liegen zwei Variablen vor, die Zeitfenster erfassen. Zunächst wird die gröbere Klassifizierung nach Jahrzehnt, in dem die Datenerhebung stattgefunden hat beziehungsweise auf das sich die Daten beziehen, betrachtet. Wenig überraschend wurde ein großer Teil der Daten in den 1990er-Jahren erhoben, eine befriedigend hohe Fallzahl liegt auch noch für die 1980er und 2000er-Jahre vor. Frühere Jahrzehnte sind aufgrund zu geringer Fallzahlen nicht analysierbar.

Tabelle 19: Deskriptive Statistik, Gruppierungsvariable „Jahrzehnt“

Jahrzehnt	Häufigkeit im Sample
1970er	1
1980er	14
1990er	102
2000er	21
Gesamt	138

Aufgrund der Kodierung der Variablen, die Unschärfe an den Rändern erzeugte (vergleiche Abschnitt 5.3.2), ist es problematisch, zwei direkt aufeinanderfolgende Zeitfenster zu vergleichen. Um eine ausreichende Trennschärfe zwischen den Gruppen zu erreichen vergleiche ich daher Datensätze, die in den 1980er erhoben wurden mit Datensätzen, die nach der Jahrtausendwende erstellt wurden. Auch hier sind die resultierenden Gruppenklassen relativ klein (14 und 21), was die Durchführung eines t-Tests erschwert. Dennoch zeigt die Teststatistik hier sehr deutliche Unterschiede zwischen den Gruppen auf. Die Mittelwerte unterscheiden sich nachhaltig (0,11 in den 1980ern zu 0,26 in den 2000ern). Das Signifikanzniveau des t-Tests lässt die Annahme der Hypothese mit einer geringen Fehlerwahrscheinlichkeit zu (Signifikanzniveau 0,064).

Tabelle 20: t-Test, Gruppierungsvariable „Jahrzehnt“

Jahrzehnt	N	Mittelwert	Standardabweichung
1980er	14	0,1119	0,21910
2000er	21	0,2550	0,20788

	T-Wert	Freiheitsgrade	Signifikanzniveau	Mittlere Differenz	Standardfehler der Differenz
Korrigierte Korrelation	-1,933	26,971	0,064*	-0,14315	0,07407

N = 35, * signifikant auf 10 % Niveau

Die zweite Variable zur Erfassung des Zeitrahmens beschreibt kürzere Phasen, jeweils einen Konjunkturzyklus. Die Phasen sind hierbei in etwa halb so lang wie in der Kodierung nach Jahrzehnten. Ein Blick auf die deskriptive Statistik zeigt auch hier relativ klare Unterschiede

zwischen frühen und späten Werten. Fraglich ist, ob diese Unterschiede statistisch signifikant ausfallen.¹²³

Tabelle 21: Deskriptive Statistik, Gruppierungsvariable „Konjunkturzyklus“

Konjunkturzyklus	Häufigkeit im Sample	Mittelwert der abh. Variablen	Standardabweichung der abh. Variablen in der Gruppe
1984-1987	6	0,1347	0,23690
1988-1991	22	0,0982	0,12928
1992-1996	48	0,2563	0,18668
1997-2003	55	0,2733	0,17001
Gesamt	131	0,2295	0,18396

N = 131

Der t-Test, der nur zwei Gruppen vergleichen kann, ist zur Überprüfung der Hypothese nicht geeignet, da die Gruppierungsvariable das Sample in vier unterscheidbare Gruppen aufteilt. Ich wende daher das ANOVA-Verfahren zur Analyse von Varianzen (**analysis of variance**) an. Dieses testet, im Gegensatz zum t-Test, auf Mittelwertgleichheit zwischen allen beteiligten Subsamples (der t-Test testet auf Unterschiedlichkeit). Ich vergleiche alle Zeitfenster, die mehr als zehn Fälle enthalten, also die Zeitfenster ab 1988. Eine univariate ANOVA vergleicht die Streuung von Varianzen innerhalb der untersuchten Fallgruppen mit der Varianz zwischen den Gruppen im Hinblick auf eine intervallskalierte abhängige Variable. Ist die Streuung zwischen den Gruppen deutlich größer als innerhalb der Gruppen, kann davon ausgegangen werden, dass die Gruppen in sich homogener sind als es die Gesamtgruppe war, sich die Mittelwerte der Teilgruppen also unterscheiden. Der Rückschluss auf die Mittelwerte ist jedoch nur zulässig, wenn die Varianzen innerhalb der Gruppen annähernd gleich sind. Eine weitere Voraussetzung der ANOVA betrifft die Skalenniveaus der abhängigen und erklärenden Variablen: Die abhängige Variable muss intervallskaliert und normalverteilt sein, ersteres ist bei den hier betrachteten Korrelationskoeffizienten der Fall, letzteres wurde bereits in Abschnitt 5.4.2 anhand des Kolmogorov-Smirnov Tests überprüft. Die erklärenden Variablen dagegen müssen kategorial skaliert sein. Im Rahmen von Metaanalysen liegen Gruppierungsvariablen vor, die naturgemäß nicht stetig sondern kategorial skaliert sind. Das ANOVA-Verfahren stellt damit also ein gutes Instrument zur Abschätzung von mit diesen Gruppierungsvariablen verbundenen Effektstärken dar.

Wie obige Tabelle zeigt unterscheiden sich die Standardabweichungen, die quadriert die Varianz ergeben, kaum, mit Ausnahme der Gruppe „1988-1991“, hier ist die Varianz deutlich geringer als in den anderen drei Gruppen. Fraglich ist nun, ob diese Abweichungen überzufällig und damit so gravierend sind, dass die ANOVA-Methodik nicht mehr anwendbar ist. Mit dem Levene-Verfahren liegt ein statistisches Instrument vor, das diese Frage beantworten kann. Es testet die Nullhypothese, dass die Varianzen in Teilgruppen gleich sind. Das Signifikanzniveau der Teststatistik ergibt sich aus dem Vergleich der resultierenden Werte des Levene-Tests mit einer Fisher-Verteilung (F-Verteilung). Gute Signifikanzwerte bedeuten, dass die Annahme der Nullhypothese, die Varianzen seien gleich, mit hoher Wahrscheinlichkeit falsch und das ANOVA-Verfahren damit nicht anwendbar wäre (Hartung et al., 2009). Wie die

¹²³ Aufgrund zu geringer Fallzahlen habe ich alle Fälle eliminiert, die vor 1984 erhoben wurden. Ebenfalls nicht eingeordnet werden können Fälle, deren Zeithorizont sich über mindestens zwei volle Konjunkturzyklen erstreckt.

folgende Tabelle zeigt, ist dies nicht der Fall. Der Signifikanzwert von 0,144 bedeutet, dass Annahme der Varianzunterschiedlichkeit nicht mit einer ausreichend hohen Wahrscheinlichkeit gehalten werden kann. Das ANOVA-Verfahren kann also Anwendung finden.

Tabelle 22: Levene-Test und ANOVA, Gruppierungsvariable „Konjunkturzyklus“

Levene-Statistik	Signifikanz
1,969	0,144

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,514	2	0,257	8,836	0,000***
Konstanter Term	4,666	1	4,666	160,356	0,000***
„Konjunkturzyklus“	0,514	2	0,257	8,836	0,000***
Fehler	3,550	122	0,029		
Gesamt	11,023	125			
Korrigierte Gesamtvariation	4,064	124			

N = 125, *** signifikant auf 1-Prozent Niveau, $R^2 = 0,127$

Das Gesamtmodell der ANOVA ist hochsignifikant. Relevant zur Analyse der zu testenden Hypothese sind insbesondere das Reliabilitätsmaß sowie Post-Hoc Tests. Im Rahmen der univariaten ANOVA wird ein Reliabilitätsmaß berechnet, das angibt, welcher Teil der Streuung der abhängigen Variablen durch das berechnete Modell erklärt wird. Dieser R^2 -Wert ist das Verhältnis der im Schätzmodell erklärten Varianz der abhängigen Variablen zur Gesamtvarianz.¹²⁴ Die hier berechnete ANOVA weist ein R^2 von 0,127 auf, das Modell erklärt also knapp 13 Prozent der Gesamtvarianz der abhängigen Variablen. Für den Hypothesentest ist von Belang, inwieweit sich die einzelnen Teilgruppen statistisch signifikant voneinander unterscheiden. Dies lässt sich im Zuge der ANOVA mit post-hoc Tests ermitteln. Diese basieren, ähnlich wie t-Tests, auf paarweisen Mittelwertvergleichen. Der Terminus post-hoc rührt daher, dass diese Tests grundsätzlich dazu dienen, ANOVA-Ergebnisse weiter zu analysieren, also die Unterschiede zwischen den Gruppen zu untersuchen, und damit erst nach der eigentlichen ANOVA (post hoc) durchgeführt werden. Verwendung findet hier der Scheffé-Test, der neben paarweisen Vergleichen auch versucht, Ausprägungen der Gruppenvariablen zu clustern. Der Scheffé-Test wird für Situationen empfohlen, in denen mehr als zwei Gruppen verglichen werden sollen (Hartung, Elpelt & Klösener, 2009).

Tabelle 23: Scheffé-Test, Gruppierungsvariable „Konjunkturzyklus“

I (Konjunkturzyklus)	J (Konjunkturzyklus)	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
1988-1991	1992-1996	-0,1581	0,04392	0,002***
	1997-2003	-0,1751	0,04303	0,000***
1992-1996	1988-1991	0,1581	0,04392	0,002***

¹²⁴ Für den Fall eines Modells mit einer erklärenden und einer erklärten Variablen entspricht der R^2 -Wert dem ebenfalls gebräuchlichen Maß zur Messung von Effektstärken, Eta-Quadrat (η^2), vergleiche neben Hartung et al. (2009) als Beispiel für eine praktische Anwendung von Effektstärkenmessgrößen auch Fan & Chen (2001).

	1997-2003	-0,0170	0,03369	0,881
1997-2003	1988-1991	0,1751	0,04303	0,000***
	1992-1996	0,0170	0,03369	0,881

N = 125, *** signifikant auf 1-Prozent Niveau

Die Betrachtung der resultierenden Signifikanzwerte zeigt, dass sich das Zeitfenster 1988-1991 von den beiden folgenden deutlich abhebt. Die beiden späteren Zeitfenster unterscheiden sich dagegen kaum..

Selektionsfilter Börsennotierung

Die dritte auf die Populationsselektion bezogene Gruppierungsvariable ist die Börsennotierung. Diese bildet lediglich zwei vergleichbare Gruppen, es kann zur Auswertung also wieder auf einen t-Test zurückgegriffen werden. Insgesamt 78 Studien lassen sich einer der beiden Gruppen zuordnen, die Gruppen sind damit groß genug für statistische Vergleiche. Die Ergebnisse sind hierbei sehr eindeutig. Die Mittelwerte (0,30 für nicht börsennotierte und 0,17 für börsennotierte Unternehmen) unterscheiden sich deutlich, der t-Wert ist mit einer Fehlerwahrscheinlichkeit von unter einem Prozentpunkt ungleich Null, die Gruppen damit überzufällig unterschiedlich. Modelliert als univariate ANOVA mit der Börsennotierung als einziger erklärender Variablen ergibt sich ein R^2 von 0,09.

Tabelle 24: Deskriptive Statistik, t-Test und univariate ANOVA, Gruppierungsvariable „Börsennotierung“

	Mittelwert	N	Standardabweichung
Nicht börsennotiert	0,2982	23	0,18535
börsennotiert	0,1716	55	0,18104

	T-Wert	Freiheitsgrade	Signifikanzniveau	Mittlere Differenz	Standardfehler der Differenz
Korrigierte Korrelation	2,771	40,433	0,008***	0,12665	0,04571

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,260	1	0,260	7,828	0,007***
Konstanter Term	3,579	1	3,579	107,689	0,000***
„Börsennotierung“	0,260	1	0,260	7,828	0,007***
Fehler	2,526	76	0,033		
Gesamt	6,190	78			
Korrigierte Gesamtvariation	2,786	77			

$R^2 = 0,093$, N = 78, *** signifikant auf 1-Prozent Niveau

Der Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen ist bei börsennotierten Unternehmen deutlich schwächer ausgeprägt als bei nicht börsennotierten.

5.4.3.2 Appropriation

H2: Die Korrelation zwischen Performance und Ressource fällt schwächer aus, je weiter sich die zu Grunde gelegte Performanceoperationalisierung vom Konzept der Kernperformance entfernt, je mehr sie also bereits Objekt des Verteilungskampfes zwischen Stakeholdergruppen war.

Hypothese H2 umfasst mehrere Ebenen. Der erste Schritt ist ein Test auf Unterschiedlichkeit zwischen Performanceoperationalisierungen. Hierbei gilt es zu prüfen, ob Leistungsoperationalisierungen überhaupt einen messbaren Einfluss auf die Studienergebnisse der Primärstudien ausüben. Fällt dieser Test positiv aus ist zu untersuchen, inwieweit die Unterschiede dem Muster entsprechen, das von der Appropriationstheorie vorhergesagt wird. Es liegen zwei Gruppierungsvariablen vor. Die erste appropriationsbezogene Gruppierungsvariable „aggregierte Leistungsoperationalisierungen“ umfasst die fünf Ausprägungen accountingbasiert, marktbasierend, prozessbasiert und expertenbasiert sowie sonstige. Aus appropriationstheoretischen Überlegungen heraus sind hohe Werte für die abhängige Variable „Korrelation“ in der Gruppe prozessbasiert, geringe in den Gruppen accounting- und marktbasierend zu erwarten (vergleiche zur ausführlichen Herleitung die Abschnitte 3.3 und 5.2). Die zweite leistungsbezogene Variable umfasst einzelne Leistungsoperationalisierungen. Auf dieser disaggregierten Basis sind hohe Werte der abhängigen Variablen insbesondere bei Leistungsoperationalisierungen zu erwarten, die nah am Konzept der Kernperformance messen, etwa Umsatz, Überleben/Ausscheiden und Innovationsleistung, geringe für Indikatoren, die den Appropriationserfolg einer bestimmten Stakeholdergruppe erfassen, etwa TSR, ROA oder Gewinn.

Der Hypothesentest erfolgt entsprechend in mehreren Schritten. Zunächst wird im Rahmen einer univariaten ANOVA der Unterschied zwischen den Teilgruppen hinsichtlich der abhängigen Variablen der Metaanalyse mit der durch die Gruppenvariable Aggregierte Leistungsoperationalisierung vorgegebenen Sampleaufteilung untersucht. Im Anschluss erfolgt eine Bewertung der Mittelwerte mit Hilfe von ANOVA Post hoc Tests. Liegen hier Ergebnisse vor, kann eine feinkörnigere Analyse mit den disaggregierten Daten erfolgen.

Ein Blick auf die deskriptive Statistik zeigt zunächst die Verteilung der Fallzahlen und die Mittelwerte in den Gruppen. Die ohnehin schwer interpretierbare Ausprägung „Sonstige“ ist mit zwei Fällen unbrauchbar dünn besetzt und wird entsprechend aus den weiteren Betrachtungen entfernt. Ein erster Blick auf die Mittelwerte lässt die Vermutung zu, dass sich Hypothese H2 bestätigen könnte.

Tabelle 25: Deskriptive Statistik und Levene-Test, Gruppierungsvariable „Leistungsoperationalisierung aggregiert“

Leistungsoperationalisierung aggregiert	Mittelwert	N	Standardabweichung
Accountingbasiert	0,1680	47	0,15816
Marktbasierend	0,1251	21	0,17638
Prozessbasiert	0,2555	35	0,18842
Expertenbasiert	0,3483	33	0,14332
Gesamt	0,2276	136	0,18324

Levene-Statistik	Signifikanz
0,420	0,739

N = 136

Auffällig ist hier, dass die Standardabweichungen in den Teilgruppen mit Ausnahme der „Sammelposition“ prozessbasiert unter der des Gesamtsamples (0,187) liegen. Es ist also eine sichtbare Homogenisierung der Teilgruppen zu beobachten. Dies war bei den bisher getesteten Gruppierungsvariablen nicht der Fall. Die relativ hohe Heterogenität in der Gruppe prozessbasiert könnte darauf hindeuten, dass die Aggregation hier nicht ideal geglückt ist.¹²⁵ Auch hier sind wieder mehrere Gruppen miteinander zu vergleichen, es ist also die Anwendung des ANOVA-Verfahrens notwendig. Voraussetzung ist wiederum Varianzgleichheit in den Teilgruppen. Die Levene-Teststatistik ergibt einen Signifikanzwert von 0,74, die Annahme der Hypothese der Varianzunterschiedlichkeit ist also mit 74-prozentiger Wahrscheinlichkeit ein Fehler, daher kann von statistischer Varianzgleichheit in den Teilgruppen ausgegangen werden. Die univariate ANOVA ist also anwendbar. Wie bereits in obiger ANOVA findet auch hier der Scheffé-Test zur Analyse der Unterschiede zwischen den Teilgruppen Verwendung. Der Scheffé-Test ermöglicht auch eine Gruppierung der Teilgruppen, dies ist hier von Relevanz, da so Zusammenhänge zwischen den Leistungsoperationalisierungen klar werden. Konzeptionell wäre zu erwarten, dass accounting- und marktbasierende Kennzahlen ähnliche Ergebnisse erbringen, während sich die anderen beiden Gruppen jeweils unterscheiden sollten.

Tabelle 26: ANOVA und Scheffé-Test für Gruppierungsvariable „Leistungsoperationalisierung aggregiert“

Univariate ANOVA

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,896	3	0,299	10,837	0,000***
Konstanter Term	6,296	1	6,296	228,482	0,000***
„Leistungsoperationalisierung aggregiert“	0,896	3	0,299	10,837	0,000***
Fehler	3,637	132	0,028		
Gesamt	11,580	136			
Kor. Gesamtvariation	4,533	135			

Post-Hoc Test Paarweiser Vergleich (Scheffé-Prozedur)

I (Leistungsoperat. aggregiert)	J (Leistungsoperat. aggregiert)	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
Accountingbasiert	Marktbasiert	0,0429	0,04357	0,809
	Prozessbasiert	-0,0875	0,03706	0,140
	Expertenbasiert	-0,1803	0,03770	0,000***
Marktbasiert	Accountingbasiert	-0,0429	0,04357	0,809

¹²⁵ Zu beachten ist hier auch, dass sich Varianz und Standardabweichung mit einer Verkleinerung des Samples in den Teilgruppen aufgrund des random sampling errors (vergleiche Hunter & Schmidt, 2007) tendenziell erhöht. So liegen etwa für Gruppe 7 nur 21 Fälle vor, hier wäre also eine höhere Heterogenität zu erwarten.

	Prozessbasiert	-0,1304	0,04582	0,048**
	Expertenbasiert	-0,2232	0,04634	0,000***
Prozessbasiert	Accountingbasiert	0,0875	0,03706	0,140
	Marktbasiert	0,1304	0,04582	0,048**
	Expertenbasiert	-0,0928	0,04028	0,156
Expertenbasiert	Accountingbasiert	0,1803	0,03770	0,000***
	Marktbasiert	0,2232	0,04634	0,000***
	Prozessbasiert	0,0928	0,04028	0,156

Scheffé-Prozedur, Gruppierungen für $\alpha = 0,05$

	1	2	3
Marktbasiert	0,1298		
Accountingbasiert	0,1697		
Prozessbasiert		0,2568	
Expertenbasiert			0,3481

N = 136, $R^2 = 0,198$, ** signifikant auf 5-Prozent Niveau, *** signifikant auf 1-Prozent Niveau

Die ANOVA- und post-hoc Tests zeigen das die Hypothese unterstützende Bild. Es liegen sehr klare Unterschiede zwischen den Gruppen vor (ANOVA-Test signifikant auf < 1 Prozent Niveau). Dem paarweisen Vergleich ist zu entnehmen, dass sich accounting- und marktbasierete Leistungsoperationalisierungen nicht signifikant voneinander unterscheiden. Prozessbasierte unterscheiden sich auf einem 5-Prozent-Niveau von marktbasiereten, nicht jedoch signifikant von den anderen beiden Gruppen, expertenbasierte Operationalisierungen sind auf einem 1-Prozent Niveau von accounting- und marktbasiereten zu unterscheiden. Die folgende Clusterung zeigt das aufgrund der bisherigen Ergebnisse zu erwartende Bild: Accounting- und marktbasierete Kennzahlen erbringen statistisch ähnliche Korrelationskoeffizienten, während sich prozessbasierte und expertenbasierte Operationalisierungen hiervon und voneinander abheben.

Dies deckt sich mit den Voraussagen der Hypothese. Dieser zufolge sollten accounting- und marktbasierete Kennzahlen, die beide die Appropriationsleistung der Eigentümer messen, relativ ähnlich sein. Die sehr hohen Korrelationen, die mit expertenbasierten Leistungsindikatoren einhergehen, waren so nicht zu erwarten, dieses Ergebnis wird im weiteren Gang der Arbeit noch genauer zu diskutieren sein (vergleiche die Abschnitte 5.5 und 6). Bezüglich der drei Indikatorengruppen accountingbasiert, marktbasieret und prozessbasiert wird Hypothese H2 also bestätigt.

Nun betrachte ich die Ebene einzelner Leistungsoperationalisierungen, teile das Sample also anhand dieser Gruppierungsvariablen. Die deskriptive Statistik zeigt, dass in vielen Fällen zu wenige Ausprägungen für robuste Berechnungen vorliegen. Hier gilt es abzuwägen zwischen der Notwendigkeit, ein breites Bild an Leistungsoperationalisierungen abbilden zu können, und der Mindestgruppengröße, die statistisch als noch akzeptabel zu werten ist. Als Kompromiss zwischen der notwendigen Breite an Operationalisierungen und statistischer Robustheit betrachte ich im Folgenden nur Operationalisierungen, zu denen wenigstens neun unterschiedliche Fälle vorliegen. Es ist klar, dass bei derart geringen Fallzahlen kaum mehr statistische Signifikanzen zu erwarten sein werden, auf deskriptiver Ebene sollten jedoch Ergebnisse

generierbar sein. Folgende Tabelle gibt die deskriptive Statistik für die entstehenden Gruppen; Leistungsoperationalisierungen, die in der Tabelle gefettet dargestellt sind, übernehme ich in die ANOVA.

Tabelle 27: Deskriptive Statistik, Gruppierungsvariable „Leistungsoperationalisierungen disaggregiert“

Leistungsoperationalisierung	N	Mittelwert
Gewinn	9	0,1600
Kapitalrentabilität	31	0,1817
Return on Sales	7	0,1173
Cumulative Abnormal Returns	1	-0,1895
Jensen's Alpha	1	0,1137
Marktkapitalisierung	3	0,2863
Market-to-Book Ratio	13	0,0923
Total Shareholder Return	3	0,2147
Ausschuss	1	0,2526
Innovationsoutput	2	0,1768
Mitarbeiterzahl	1	0,2905
Produktivität (Umsatz pro Mitarbeiter)	12	0,2642
Überleben/Ausscheiden	2	0,2716
Umsatz	17	0,2548
Interne Experten	33	0,3483
Andere Operationalisierungen	2	0,3537
Gesamt	138	0,2295

Sechs Leistungsoperationalisierungen können also einer tieferen Analyse unterzogen werden. Zunächst ist erneut ein Levene-Test auf Varianzgleichheit in den Teilgruppen nötig, um zu überprüfen, ob das Scheffé-Verfahren als Post-Hoc Test anwendbar ist. Wie die folgende Tabelle zeigt ist dies der Fall.

Tabelle 28: Levene-Test, Gruppierungsvariable „Leistungsoperationalisierungen disaggregiert“

Levene-Statistik	Signifikanz
1,059	0,387

N = 115

Klar ist auch, dass bei den vorliegenden sehr kleinen Gruppen Signifikanztests, die wesentlich auf der einbezogenen Fallzahl basieren, nur mehr bei sehr stark ausgeprägten Unterschieden Ergebnisse erbringen können. Dies zeigt sich insbesondere bei den sehr kleinen Gruppen Ge-

winn, Market-to-Book Ratio und Produktivität, die sich in den paarweisen Vergleichen zwar in absoluten Zahlen deutlich von den anderen Operationalisierungen unterscheiden, diese Unterschiede sind aufgrund der geringen Zahl der Beobachtungen kaum mehr statistisch signifikant.

Tabelle 29: ANOVA und Scheffé-Test, Gruppierungsvariable „Leistungsoperationalisierung disaggregiert“

Univariate ANOVA

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,843	5	0,169	5,916	0,000***
Konstanter Term	4,312	1	4,312	151,222	0,000***
„Leistungsoperationalisierung disaggregiert“	0,843	5	0,169	5,916	0,000***
Fehler	3,108	109	0,029		
Gesamt	10,418	115			
Korrigierte Gesamtvariation	3,952	114			

N = 115, R² = 0,213

Post-Hoc Test (Scheffé-Prozedur)

I (Leistungsoperationalisierung)	J (Leistungsoperationalisierung)	Mittlere Differenz (I-J)	Standardfehler	Signifikanz
Gewinn	Kapitalrentabilität	-0,0217	0,06394	1,000
	Market-to-Book Ratio	0,0677	0,07322	0,973
	Produktivität	-0,1042	0,07446	0,854
	Umsatz	-0,0949	0,06961	0,867
	Interne Experten	-0,1883	0,06350	0,127
Kapitalrentabilität	Gewinn	0,0217	0,06394	1,000
	Market-to-Book Ratio	0,0894	0,05580	0,765
	Produktivität	-0,0825	0,05741	0,839
	Umsatz	-0,0731	0,05096	0,840
	Interne Experten	-,01666*	0,04224	0,012***
Market-to-Book Ratio	Gewinn	-0,0677	0,07322	0,973
	Kapitalrentabilität	-0,0894	0,05580	0,765
	Produktivität	-0,1719	0,06760	0,272
	Umsatz	-0,1625	0,06222	0,243
	Interne Experten	-,02560*	0,05530	0,001***
Produktivität	Gewinn	0,1042	0,07446	0,854
	Kapitalrentabilität	0,0825	0,05741	0,839
	Market-to-Book Ratio	0,1719	0,06760	0,272

	Umsatz	0,0093	0,06367	1,000
	Interne Experten	-0,0841	0,05692	0,822
Umsatz	Gewinn	0,0949	0,06961	0,867
	Kapitalrentabilität	0,0731	0,05096	0,840
	Market-to-Book Ratio	0,1625	0,06222	0,243
	Produktivität	-0,0093	0,06367	1,000
	Interne Experten	-0,0935	0,05041	0,634
Interne Experten	Gewinn	0,1883	0,06350	0,127
	Kapitalrentabilität	,01666*	0,04224	0,012***
	Market-to-Book Ratio	,02560*	0,05530	0,001***
	Produktivität	0,0841	0,05692	0,822
	Umsatz	0,0935	0,05041	0,634

N = 115, *** signifikant auf 1-Prozent Niveau

Die Betrachtung der individuellen Leistungsoperationalisierungen bestätigt im Wesentlichen das oben gewonnene Bild: Accounting- und marktbasierende Operationalisierungen messen die deutlich geringsten Korrelationen zwischen Leistung und Ressourcen, wobei marktbasierende Indikatoren noch deutlich geringere Beziehungen anzeigen als accountingbasierte. Expertenbasierte wiederum bilden klar und statistisch robust die stärksten Korrelationen ab. Aufgrund der geringen Fallzahlen in den Gruppen sind die Ergebnisse statistisch nicht überzubewerten, bieten aber dennoch, grafisch abgetragen, zumindest auf einer deskriptiven Ebene das die Hypothese H2 bestätigende Bild:

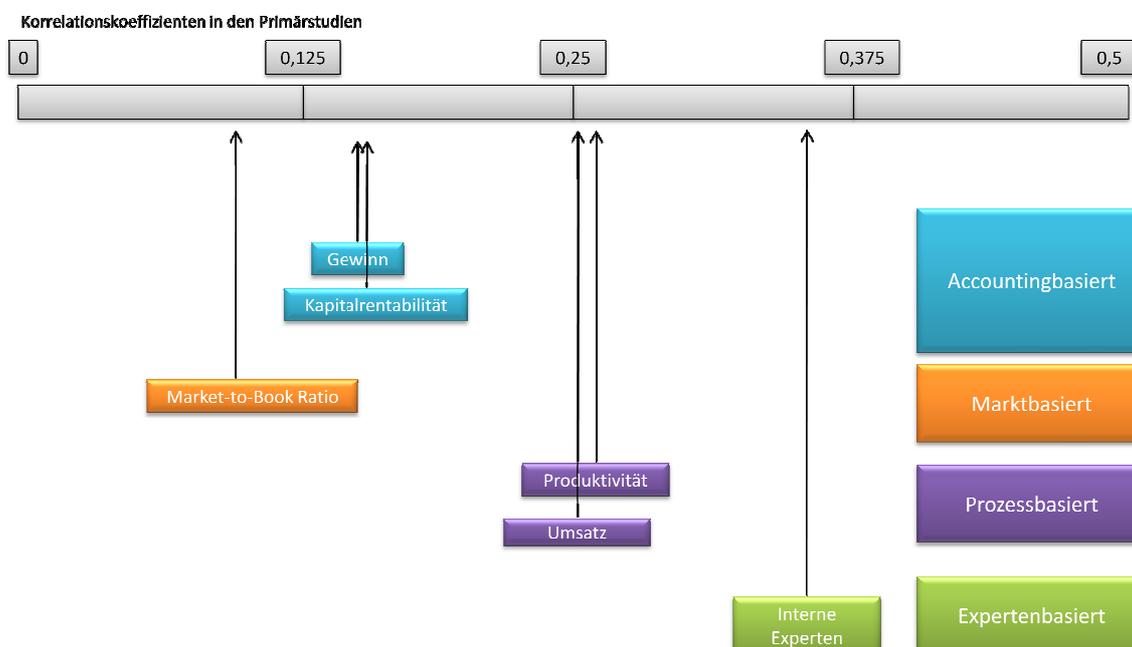


Abbildung 6: Leistungsoperationalisierungen disaggregiert, deskriptiv

Primärstudien, die mit Leistungsoperationalisierungen arbeiten die die Leistungsappropriation der Anteilseigner messen (Gewinn, Kapitalrentabilität, Market-to-Book Ratio), berichten deutlich geringere Korrelationen zwischen Performance und Ressourcen im Sinne des Resource-based view als solche, die die Leistung nahe am Konzept der Kernperformance erfassen.

sen (Produktivität, Umsatz). Die Annahme, dass die relative Verhandlungsmacht von Stakeholdergruppen als unbeobachtete Moderatorvariable die Aufdeckung von Beziehungen zwischen abhängigen und unabhängigen Variablen in der Strategischen Managementforschung erschweren könnte wird von den empirischen Ergebnissen gestützt.

Betrachtet man abschließend nun noch die R^2 -Werte, also die Varianzaufklärung der gerechneten Modelle, ergibt sich ein R^2 -Werte von 0,198 für die Modellierung der univariaten ANOVA mit der erklärenden Variablen „aggregierte Leistungsoperationalisierungen“. Der R^2 -Wert für die Betrachtung disaggregierter Leistungsoperationalisierungen liegt bei 0,213. Die Gruppe, zu der eine in einer Primärstudie angewandte Leistungsoperationalisierung gehört, erklärt also rund 20 Prozent der Varianz der Ergebnisse der betrachteten Primärstudien.

5.4.3.3 Zeitversatzproblem

H3.1: Studien, die mit für Zeitversatzprobleme robusten Leistungsoperationalisierungen arbeiten, finden einen stärkeren Zusammenhang zwischen Ressourcen und Performance als Studien, deren Performanceoperationalisierungen nicht robust gegenüber Zeitversatzphänomenen sind.

H3.2: Der Teil der Studien, der nicht mit für Zeitversatz robusten Verfahren arbeitet, weist eine höhere Varianz in der abhängigen Variablen auf als der Teil, der robuste Verfahren einsetzt.

Im Rahmen der Tests für den Einfluss von Time-Lags auf Studienergebnisse ist zunächst, wie auch in den anderen Fällen, auf Unterschiedlichkeit der Teilsamples zu kontrollieren. Hier liegt eine Gruppierungsvariable vor, die das Sample der Primärstudien in lediglich zwei Gruppen teilt, je nachdem, ob mit für Zeitversatzprobleme robusten Verfahren gearbeitet wurde oder nicht. Die deskriptive Statistik gibt berechtigten Grund zu der Annahme, dass sich die beiden Gruppen unterscheiden könnten. Alle 138 in der Metaanalyse betrachten Performance-Ressourcen-Beziehungen können eindeutig einer der beiden Gruppen zugeordnet werden, die resultierenden Gruppen sind also umfangreich genug für weitere Analyse. Der Mittelwert der Gruppe, die auf für Time-Lag robusten Leistungsoperationalisierungen basiert, ist deutlich höher als der Mittelwert der Korrelationskoeffizienten der Vergleichsgruppe (0,27 zu 0,19). Die Varianzen unterscheiden sich zwischen den beiden Gruppen kaum und liegen in etwa auf dem Niveau des Gesamtsamples.

Tabelle 30: Deskriptive Statistik, Gruppierungsvariable „Zeitversatz“

Robust für Zeitversatzprobleme?	Mittelwert	N	Standardabweichung
Nein	0,1851	63	0,18761
Ja	0,2668	75	0,17341
Gesamt	0,2295	138	0,18396

Die Ermittlung der statistischen Signifikanz der Mittelwertunterschiede kann hier, da nur zwei Ausprägungen der erklärenden Variablen vorliegen, wieder mit einem t-Test erfolgen. Der Test auf Gleichheit der Varianzen erfolgt mit Hilfe des Levene-Tests, der R^2 -Wert wird auf Basis einer univariaten ANOVA berechnet, wenn der Levene-Test nicht signifikant ausfallen

sollte. Der t-Test erbringt das erwartete Resultat: Der Unterschied in den Mittelwerten ist statistisch hoch signifikant. Hypothese H3.1 wird also bestätigt, Studien, die mit für Zeitversatzprobleme robusten Verfahren zur Leistungsmessung arbeiten, finden statistisch belastbar stärkere Zusammenhänge zwischen der Leistung von Organisationen und Ressourcen im Sinne des Resource-based view als solche, die dies nicht tun.

Tabelle 31: t-Test und Levene-Test, Gruppierungsvariable „Zeitversatz“

Levene-Statistik	Signifikanz
0,110	0,740

T-Test

	T-Wert	Freiheitsgrade	Signifikanzniveau	Mittlere Differenz	Standardfehler der Differenz
Korrigierte Korrelation	-2,638	136	0,009***	-0,08171	0,03077

Univariate ANOVA

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,229	1	0,229	7,054	0,009***
Konstanter Term	6,990	1	6,990	215,674	0,000***
„Robust gegen Zeitversatzproblem“	0,229	1	0,229	7,054	0,009***
Fehler	4,408	136	0,032		
Gesamt	11,902	138			
Korrigierte Gesamtvariation	4,636	137			

N = 138, $R^2 = 0,049$, *** signifikant auf 1-Prozent Niveau

Hypothese H3.2 lässt sich mit Hilfe des Levene-Tests auf Varianzgleichheit bewerten. Der Blick auf die deskriptive Statistik lässt es bereits erahnen, statistisch belastbare Unterschiede in den Varianzen zwischen den beiden Gruppen sind nicht zu beobachten. Der Levene-Test auf Varianzgleichheit ist nicht signifikant, die Annahme der Hypothese, dass sich die Varianzen der beiden Teilgruppen unterscheiden, wäre mit 74-prozentiger Wahrscheinlichkeit ein Fehler, daher kann Hypothese H3.2 nicht angenommen werden. Die univariate ANOVA ergibt einen R^2 -Wert von 0,049, der Time-Lag erklärt also knapp fünf Prozent der Varianz der abhängigen Variablen.

Zu beachten ist allerdings, dass diese Auswertung in der Kodierung nicht unabhängig von der Kodierung und damit den Ergebnissen der Appropriationsproblematik ist. Die Erfassung der Leistung von Organisationen durch die Befragung von Experten wurde als eine Lösungsmöglichkeit zur Bewältigung des Zeitversatzproblems angeführt. Die Untersuchung in Abschnitt 5.4.2 zeigte jedoch, dass die Operationalisierung der Leistung mittels Expertenbefragung grundsätzlich in deutlich höheren Korrelationen zwischen Leistung und Ressourcen resultiert

als dies bei anderen Leistungsoperationalisierungen der Fall ist. Unklar ist nun, worauf diese hohe Korrelation zurückzuführen ist. Die Betrachtung des Zeitversatzproblems könnte ein möglicher Ansatzpunkt für eine Erklärung sein: Experten könnten deshalb die Zusammenhänge zwischen einer bestimmten Ressource und der Leistung einer Organisation besser abbilden können, da sie in der Lage sind, zukünftige Entwicklungen näherungsweise befriedigend zu prognostizieren, also das Zeitversatzproblem zu umgehen. Es ist aber auch möglich, dass die Operationalisierung der Leistung über Experten aus anderen Gründen hohe Korrelationen erbringt. Dies wird im weiteren Gang der Arbeit noch detaillierter zu diskutieren sein.

Zunächst einmal ist es wichtig zu überprüfen, ob die Leistungsoperationalisierung „Experte“ das Zeitversatzproblem beeinflusst, ob also die getroffenen Aussagen zum Zeitversatzproblem nur für den spezifischen Fall der Experten gelten oder ob auch die anderen, konzeptionell aus der Diskussion des Zeitversatzproblems (Abschnitt 3.4.2) auf stabilerem Fundament stehenden Lösungsmöglichkeiten, wie etwa die Nutzung zeitlich versetzter abhängiger Variablen, zu einer Erhöhung der in den Primärstudien beobachteten Korrelationen zwischen Performance und Ressource führen. Um dies zu überprüfen, berechne ich noch einmal eine ANOVA, mit einem Sample aus dem die expertenbasierten Leistungsoperationalisierungen eliminiert sind. Die Mittelwerte zwischen den Gruppen unterscheiden sich nach Eliminierung der Primärstudien, die mit Experteneinschätzungen arbeiten, nur mehr wenig (0,18 bei nicht timelagrobusten zu 0,21 bei timelagrobusten Verfahren). Die folgende Tabelle zeigt das Ergebnis der ANOVA mit verkleinertem Sample:

Tabelle 32: ANOVA, Sample ohne expertenbasierte Leistungsoperationalisierungen, Gruppierungsvariable „Zeitversatz“

Univariate ANOVA

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,023	1	0,023	0,701	0,405
Konstanter Term	3,898	1	3,898	120,070	0,000***
„Robust gegen Zeitversatzproblem“	0,023	1	0,023	0,701	0,405
Fehler	3,344	103	0,032		
Gesamt	7,242	105			
Korrigierte Gesamtvariation	3,366	104			

N = 105, 57 davon timelagrobust, 48 nicht timelagrobust, $R^2 = 0,007$, *** signifikant auf 1-Prozent Niveau

Wie der Tabelle zu entnehmen ist verliert das Modell nun seine statistische Signifikanz und erklärt praktisch keine Varianz mehr. Der Unterschied in den Mittelwerten ist nicht mehr als überzufällig, das Zeitversatzproblem nicht mehr als statistisch signifikant anzusehen, wenn die hohen Korrelationen, die mit der Operationalisierung über Experten einhergehen, aus der Rechnung herausgenommen werden. Der Grund hierfür wird im Weiteren näher zu beleuchten sein.

5.4.4 Weiterführende Untersuchungen

Nun sind weiterführende Untersuchungen anzustellen um die bisher gewonnenen Erkenntnisse weiter zu durchleuchten. Zunächst gilt es dem Verdacht nachzugehen, dass sich die beobachteten Einflussfaktoren auf die Performance-Ressourcen-Beziehung durch die Heterogenität des Konstrukts Ressource erklären lassen könnten; im Anschluss sind Interaktionsbeziehungen zwischen den ermittelten Gruppierungsvariablen zu betrachten.

5.4.4.1 Erklärt die Variable „Ressource“ die bisherigen Ergebnisse?

Es ist gut vorstellbar, dass bestimmte Fragestellungen in der Beziehung zwischen Leistung und Ressource regelmäßig mit bestimmten Leistungsoperationalisierungen erfasst werden und die oben aufgedeckten Unterschiede damit nicht der jeweiligen Leistungsoperationalisierung sondern der mit ihr hauptsächlich bearbeiteten Ressourcen-Performance-Beziehung geschuldet sind. So fiel im Rahmen der Literaturreview beispielsweise auf, dass die Forschung zu den Erfolgsdeterminanten von Mergern wesentlich mit CAR als Leistungsoperationalisierung arbeitet. Wäre im hier zu untersuchenden Fall etwa zu beobachten, dass besonders leistungsnah oder -ferne Ressourcen (etwa Innovationsfähigkeit oder bestimmte Personalmanagementmaßnahmen) stark bevorzugt mit einer bestimmten Leistungsoperationalisierung erfasst werden, könnte dies ein alternatives Erklärungsmuster für die beobachteten Unterschiede zwischen Performanceoperationalisierungen darstellen. Würden beispielsweise Ressourcen, die als sehr leistungsnah einzuschätzen sind und damit konzeptionell bereits eine stärkere Korrelation zur Unternehmensleistung erwarten lassen, bevorzugt mit prozessbasierten Leistungsoperationalisierungen gemessen werden und Ressourcen, die konzeptionell nicht mit der Leistung verknüpft sind, mit accounting- oder marktbasierter, würde dies die beobachteten Zusammenhänge zwischen Leistungsoperationalisierung und Korrelation zwischen Ressource und Performance ebenso gut erklären wie die angestellten appropriations-theoretischen Überlegungen.

Die Heterogenität der Ressourcen, die in der vorliegenden Metaanalyse untersucht wurde, erlaubt zumindest eine explorative Betrachtung dieser Problematik. Eine Ressource, die Verwendung von High Performance Worksystems (HPWS, 25 der 138 Fälle, 18 Prozent des Gesamtsamples), ist so häufig im Sample vertreten, dass vergleichende Untersuchungen nach Leistungsoperationalisierung innerhalb dieser Gruppen möglich sind. High Performance Work Systems sind konzeptionell nicht allzu stark mit der Leistung einer Organisation verbunden. Sie gehen, vereinfacht zusammengefasst, davon aus, dass sich bestimmte Investitionen in die Mitarbeiter für Unternehmen lohnen. Diese Investitionen sind zunächst einmal mit Kosten verbunden, die teilweise auf die Appropriationsleistung der Arbeitnehmer bezogen werden können (beispielsweise höheres Gehalt), teilweise auch Kosten in Bezug auf die Kernperformance darstellen und diese schmälern können (etwa hoher Aufwand im Mitarbeiterauswahlprozess; vergleiche auch Huselid, 1995, Combs et al., 2006). Ob sich HPWS unter dem Strich für die Kernperformance oder für die Stakeholdergruppe der Anteilseigner lohnen ist konzeptionell nicht eindeutig geklärt. Die folgende Tabelle fasst die deskriptive Statistik zu dieser Fragestellung zusammen, soweit sie aus dem Sample der Metaanalyse ableitbar ist.

Tabelle 33: Deskriptive Statistik, reduziertes Sample Ressource HPWS, Gruppierungsvariable „Leistungsoperationalisierung aggregiert“

Leistungsoperationalisierung	N	Mittelwert	Standardabweichung
Accountingbasiert	6	0,0379	0,10924
Marktbasiert	2	0,1453	0,06252
Prozessbasiert	9	0,1474	0,18756
Expertenbasiert	8	0,3347	0,14354
Gesamt	25	0,1809	0,18415

Es ist klar zu erkennen, dass die Ressource HPWS mit Leistungsoperationalisierungen aller vier Kategorien untersucht wird, eine dominante Leistungsoperationalisierung gibt es nicht. Insgesamt lässt sich ein Zusammenhang zwischen HPWS und Leistung durchaus bestätigen (Mittelwert der Korrelationen 0,18), allerdings besteht praktisch kein Zusammenhang zwischen HPWS und accountingbasierten Kennzahlen. Für die Anteilseigner scheint die Einführung von HPWS eher einem Nullsummenspiel zu gleichen, während die Kernperformance, die mit prozessbasierten Leistungsoperationalisierungen assoziiert wird, offenbar gesteigert wird. Überdies ist festzustellen, dass auch hier expertenbasierte Leistungsoperationalisierungen deutlich höhere Korrelationen zwischen HPWS und der Leistung von Organisationen anzeigen als andere Methoden der Leistungsmessung. Der Wert für marktbasierete Kennzahlen ist bei nur zwei Fällen nicht interpretierbar. Die sehr geringe Korrelation, die bei accountingbasierten Kennzahlen auftritt, fügt sich damit in das von der Appropriationstheorie gezeichnete Bild. Die Einführung von HPWS führt tendenziell zu einer Erhöhung der relativen Verhandlungsmacht der Stakeholdergruppe Mitarbeiter, accountingbasierte Kennzahlen, die die Aneignungsleistung der Anteilseigner messen, weisen kaum Zusammenhänge zwischen HPWS und dieser Performanceindikation aus. Bei den Standardabweichungen ist wieder zu beobachten, dass die Teilgruppen homogener sind als das Gesamtsample, auch dies deckt sich mit den Ergebnissen, die in Abschnitt 5.4.3.2 erzielt wurden. Auch hier fällt auf, dass die Standardabweichung in der Gruppe Prozessbasiert höher ausfällt als in den anderen drei Gruppen.

Es lässt sich also festhalten, dass die Ressource HPWS die bisher erlangten Ergebnisse nicht erklären kann, vielmehr bestätigt sich noch einmal die Argumentation, die in Abschnitt 3.3 entwickelt wurde. Aufgrund der kleinen Fallzahlen bieten sich über deskriptive Betrachtungen hinausgehende statistische Analysen nicht an.

5.4.4.2 Interaktionseffekte

Es ist denkbar, dass sich die drei untersuchten Methodenprobleme (Selektionsproblem, Appropriationsproblem und Time-Lag Problem) gegenseitig beeinflussen. Eine ANOVA mit mehreren erklärenden Variablen ist geeignet, um etwaige Interdependenzeffekte zwischen diesen erklärenden Variablen aufzuspüren. Der Interaktionseffekt zwischen zwei oder mehr erklärenden Variablen wird hierbei erfasst indem die multiplikative Verknüpfung der interessierenden Variablen zusätzlich in das Modell eingebracht wird. Hilfreich zur weiteren Analyse ist insbesondere die Erstellung einer Matrix (mehrdimensionale Kreuztabelle), in der die Effekte einer erklärenden Variablen in Abhängigkeit weiterer erfasst werden. Auf diese Weise sind Mittelwert- und Varianzvergleiche zwischen Untergruppen möglich, es ist also beispielsweise überprüfbar, inwieweit sich die Verwendung bestimmter Leistungsoperationalisierungen in

unterschiedlichen Zeitfenstern unterscheidet oder ob Studien, die mit börsennotierten Unternehmen arbeiten, häufiger oder weniger häufig zeitversatzrobuste Verfahren anwenden.

Ich bilde im Folgenden gesättigte ANOVA-Modelle¹²⁶ und vergleiche paarweise die vier bisher als statistisch signifikant ermittelten Moderatorvariablen Jahrzehnt, Notierung, Leistungsoperationalisierung (aggregiert) sowie Time-Lag-Robustheit¹²⁷. Auf diese Weise ermittle ich Interaktionseffekte erster Ordnung. Um zu wenig gefüllte Gruppen und nicht interpretierbare Verknüpfungen zu vermeiden muss der Datensatz noch angepasst werden:

- Im Rahmen der Betrachtung der Variable Jahrzehnt eliminiere ich wieder alle Datensätze, die vor 1980 erhoben wurden.
- Die Ausprägung „gemischt/keine Angabe“ der Variable Notierung wird nicht betrachtet.
- Die Leistungsoperationalisierung „Sonstige“ ist ebenfalls nicht interpretierbar, entsprechende Fälle werden eliminiert, wenn die Gruppierungsvariable „Leistungsoperationalisierung“ betrachtet wird.

Diese Anpassungen werden nur dann vorgenommen, wenn die entsprechende Variable tatsächlich in ein Modell einbezogen wird, um einen möglichst großen Datensatz in jedem einzelnen Untersuchungsschritt zu realisieren. Dennoch ist zu beachten, dass sich auch hier die Gruppengrößen teilweise am unteren Rand des statistisch sinnvollen bewegen. Im Folgenden werden Rechnungen, Ergebnisse und erste Einordnungen vorgestellt, eine ausführlichere Diskussion der Ergebnisse erfolgt im anschließenden Abschnitt 5.5.

Interaktion zwischen den Gruppierungsvariablen Jahrzehnt und Notierung: Die Relevanz der Interaktionsbeziehung zwischen Zeitraum und Notierung zeigt sich im gesättigten Modell sehr deutlich. Wird die Interaktionsbeziehung (multiplikative Verknüpfung) zwischen Notierung und Zeitraum in eine Schätzgleichung einbezogen verliert die alleinstehende Variable Jahrzehnt ihre Signifikanz auf dem 10-Prozent Niveau, der Interaktionsterm dagegen ist hoch signifikant. Die Schätzgleichung erklärt insgesamt 13 Prozent der beobachteten Varianz der abhängigen Variablen.¹²⁸

Tabelle 34: Levene-Test, ANOVA und Interaktionsmatrix Notierung / Zeitraum

Levene-Statistik	Signifikanz
1,130	0,353

¹²⁶ Ein Modell ist gesättigt, wenn alle möglichen Verknüpfungen zwischen einbezogenen erklärenden Variablen in die Modellierung eingebracht werden. Im Falle der ANOVA sind diese Verknüpfungen, wie oben erwähnt, multiplikativ.

¹²⁷ Die Gruppierungsvariable Timelag-Robustheit ist nur signifikant, sofern die Leistungsoperationalisierung mittels Experten aus dem Sample nicht entfernt wird, vergleiche Abschnitt 5.4.3.3.

¹²⁸ Bei Verwendung mehrerer erklärender Variablen empfiehlt sich die für die Gruppenzahl korrigierte Variante des R^2 -Wertes, vergleiche ausführlich Hartung et al. (2009) und Brosius (2008), auf die im Folgenden ausschließlich Bezug genommen wird.

Interaktionsmatrix Notierung und Jahrzehnt

Börsennotierung	Jahrzehnt	Mittelwert	Standardfehler
Nicht börsennotiert	1980er	0,543 (1)	0,178
	1990er	0,316 (18)	0,042
	2000er	0,155 (4)	0,089
börsennotiert	1980er	0,102 (10)	0,056
	1990er	0,168 (36)	0,030
	2000er	0,249 (8)	0,063

N = 77, in Klammern die Fallzahlen in den Gruppen

Zwischensubjekteffekte Notierung und Jahrzehnt

	Quadratsumme	Freiheitsgrade	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	0,518	5	0,104	3,282	0,010***
Konstanter Term	1,510	1	1,510	47,821	0,000***
„Notierung“	0,157	1	0,157	4,968	0,029**
„Jahrzehnt“	0,040	2	0,020	0,640	0,530
„Notierung“ * „Jahrzehnt“	0,226	2	0,113	3,578	0,033**
Fehler	2,242	71	0,032		
Gesamt	6,055	77			
Korrigierte Gesamtvariation	2,761	76			

N = 77, korrigiertes $R^2 = 0,131$, ** signifikant auf 5-Prozent Niveau, *** signifikant auf 1-Prozent Niveau

Dieses Ergebnis erklärt sich aus der Betrachtung der Interaktionsmatrix und der entsprechenden Visualisierung in unten stehender Abbildung. In den beiden Gruppen „börsennotiert“ beziehungsweise „nicht börsennotiert“ liegen zwei gegenläufige Tendenzen vor. Während über die betrachteten Zeiträume die beobachteten Korrelationen zwischen Ressourcen und Performance bei börsennotierten Organisationen tendenziell immer größer werden, ist bei nicht börsennotierten Organisationen ein gegenläufiger Effekt zu beobachten.

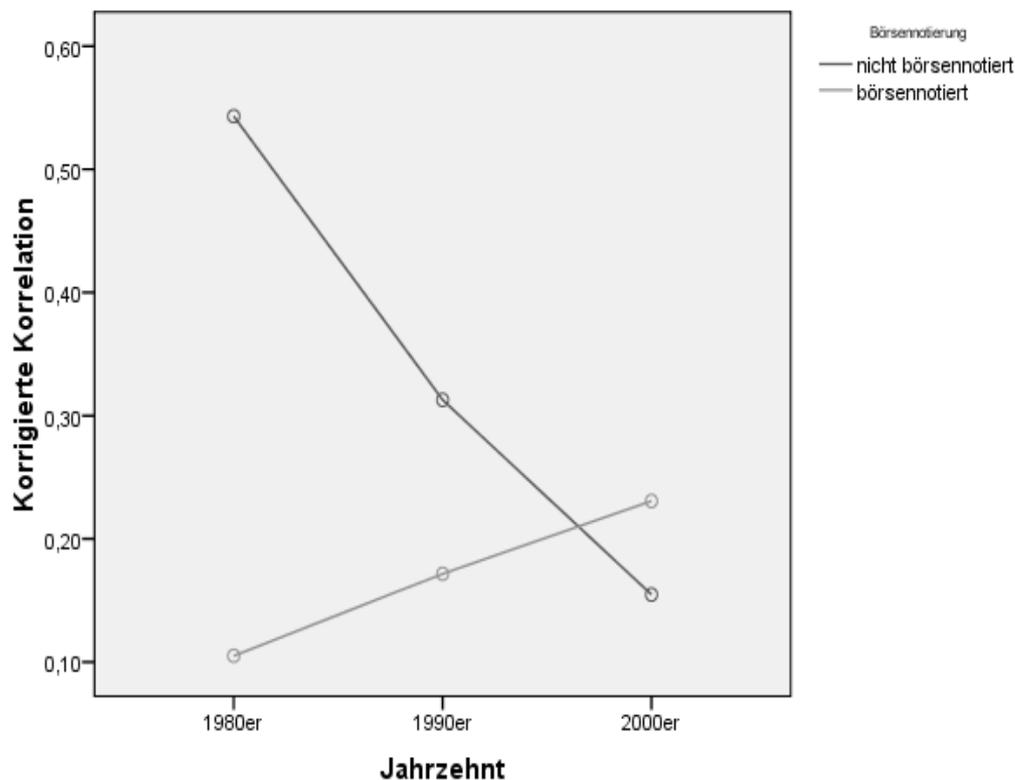


Abbildung 7: Interdependenz Jahrzehnt / Notierung

Konzeptionell liegt kein Ansatz vor, der diese überzufällige gegenläufige Tendenz erklären würde, die Beobachtung ist somit zunächst nicht ohne Rückgriff auf nicht begründbare Spekulationen erklärbar. Nicht überinterpretiert werden darf die Tendenz bei den nicht börsennotierten Unternehmen, hier sind die Gruppen in den 1980ern und 2000ern sehr klein besetzt. Die Tendenz bei den börsennotierten Unternehmen jedoch basiert auf relativ gut gefüllten Gruppen und ist daher nicht so leicht abzutun. In den Bereich der Spekulation fallende Erklärungsansätze könnten etwa dahin gehen, dass börsennotierte Unternehmen durch die zunehmende Verbreitung des Resource-based view in den letzten Jahrzehnten gelernt haben, ihre Ressourcen besser in Leistung umzusetzen. Diese Erklärung würde sich auch mit der Beobachtung im Gesamtsample decken, hier konnte in Abschnitt 5.4.3.1 ja gezeigt werden, dass in den 2000ern allgemein deutlich stärkere Korrelationen zwischen Performance und Ressourcen berichtet wurden als dies in den 1980ern der Fall war. Allerdings sind auch zahlreiche andere, ebenso spekulative Erklärungen für diese Beobachtung denkbar.

Interaktion zwischen den Gruppierungsvariablen Notierung und Leistungsoperationalisierung: Ob die Art der Leistungsoperationalisierung die Signifikanz der Gruppierungsvariable Notierung erklärt ist nicht direkt feststellbar, da die Gruppengrößen einer Interaktionsmatrix naturgemäß in einigen Feldern sehr klein beziehungsweise null werden. Allgemein fokussieren sich Studien hauptsächlich dann auf börsennotierte Unternehmen, wenn accounting- oder marktbasierende Leistungsgrößen im Fokus der Betrachtung stehen. Insgesamt beziehen sich 76

Prozent aller Beobachtungen in dieser Gruppe auf accounting- und marktbasierende Leistungsoperationalisierungen. Hieraus resultieren verschiedene kaum oder nicht gefüllte Felder der Interaktionsmatrix (nicht börsennotiert/marktbasiert, börsennotiert/expertenbasiert, nicht börsennotiert/accountingbasiert), damit ist eine ANOVA nicht mehr berechenbar.

Interaktion zwischen den Gruppierungsvariablen Jahrzehnt und Leistungsoperationalisierung: Eine Betrachtung dieses Zusammenhangs kann, neben Mittelwertunterschieden zwischen den Operationalisierungs- und Jahresgruppen, auch die Verwendung von Leistungsoperationalisierungen in den jeweiligen Zeiträumen beleuchten und ist daher auch konzeptionell interessant. Wie unten stehender Tabelle zu entnehmen ist, ist der Interaktionsterm jedoch nicht signifikant. Zudem verliert die Gruppierungsvariable Jahrzehnt im gesättigten Modell ihre Signifikanz.

Tabelle 35: ANOVA Zeit / Leistungsoperationalisierung

	Quadrat-summe	Freiheits-grade	Mittel-der	F	Signifi-kanz
Korrigiertes Modell	1,164	11	0,106	3,88	0,000***
Konstanter Term	1,700	1	1,700	62,4	0,000***
„Jahrzehnt“	0,075	2	0,038	1,37	0,256
„Leistungsoperationalisierungen aggregiert“	0,212	3	0,071	2,59 3	0,056*
„Leistungsoperationalisierungen aggregiert“ * „Jahr-“	0,086	6	0,014	0,52 7	0,787
Fehler	3,350	123	0,027		
Gesamt	11,446	135			
Korrigierte Gesamtvariation	4,514	134			

korrigiertes $R^2 = 0,191$, $N = 135$

Damit ist festzuhalten, dass die Leistungsoperationalisierung auch einen guten Teil des Einflusses der Gruppierungsvariablen Jahrzehnt erklärt. Betrachtet man die Häufigkeit der Nutzung bestimmter Leistungsoperationalisierungen im Zeitablauf fällt auf, dass in Arbeiten mit neueren Datensätzen anteilig mehr prozessbasierte Kennzahlen Verwendung finden, während der Anteil der marktbasierenden Kennzahlen von 39 Prozent in den 1980ern auf 5 Prozent in den 2000er-Jahren zurückgegangen ist. In allen drei Zeitfenstern nutzen konstant etwa ein Drittel der Studien accountingbasierte Kennzahlen. Die Leistungsoperationalisierung überlagert also den Einfluss der Gruppierungsvariablen „Jahrzehnt“.

Auf einer deskriptiven Basis lassen sich dennoch Indizien dafür finden, dass der Effekt des betrachteten Zeitraums relevant sein könnte. Accounting- und marktbasierende Operationalisierungen zeigen über die drei betrachteten Jahrzehnte zunehmend stärker werdende Korrelationen, bei prozess- und expertenbasierten Operationalisierungen ist dagegen kaum ein Einfluss zu erkennen. Die Cluster sind für die 1980er und 2000er Jahre allerdings durchgehend mit $N < 10$ zu dünn besetzt um bewerten zu können, ob dieser Effekt überzufällig stark ausgeprägt ist.

Interaktion zwischen den Gruppierungsvariablen Jahrzehnt und Zeitversatzrobustheit: Interessant ist hier insbesondere die Interaktionsmatrix die angibt inwieweit sich in unterschiedlichen Zeitfenstern der Umgang mit dem Zeitversatzproblem geändert hat. Wie die Interaktionsmatrix zeigt, wurde in den unterschiedlichen Zeitfenstern tatsächlich unterschiedlich häufig mit zeitversatzrobusten Verfahren gearbeitet. In den 1990er Jahren arbeiten deutlich mehr Studien mit Verfahren, die für Zeitversatzprobleme robust sind, als dies in den 1980er und 2000er-Jahren der Fall ist. Hier sind allerdings wieder die relativ geringen Fallzahlen für die 1980er und 2000er-Jahre zu beachten, die Vollerhebung der aktuellsten Forschung, die in Kapitel 4 dieser Arbeit besprochen wurde, zeigte ja auch ein deutlich anderes Bild, hier arbeitete die deutliche Mehrzahl der aktuellen Forschungen mit zeitversatzrobusten Verfahren.

Allerdings hängt auch die Wirkung der Time-Lag-Robustheit vom Jahrzehnt ab. In den 1980er ist der Zusammenhang so, wie Hypothese H1.4 dies voraussagt, Studien, die mit für Zeitversatzprobleme robusten Kennzahlen arbeiten, zeigen deutlich höhere Korrelationen an als solche, die dies nicht tun. In den 1990ern gibt es keinen Unterschied, wie auch in Abschnitt 5.4.3.3 errechnet. In den 2000er-Jahren kehrt sich der Zusammenhang sogar um. Konzeptionelle oder anderweitig plausible Gründe für diese Entwicklung sind nicht augenfällig, es mag sich bei den Werten für die 1980er und 2000er Jahre um einen Zufallsausschlag bedingt durch die kleinen Fallgruppen handeln. Dies spricht dafür, dass die Zeitversatzrobustheit allenfalls einen sehr kleinen Teil der Heterogenität der Studienergebnisse in der Strategischen Managementforschung erklären kann.

Tabelle 36: Interaktionsmatrix Zeit / Time Lag

Interaktionsmatrix

Jahrzehnt	Robust für Zeitversatzproblem	Mittelwert	Standardabweichung	N
1980er	Nein	-0,013	0,16987	8
	Ja	0,218	0,12562	4
1990er	Nein	0,182	0,16244	36
	Ja	0,199	0,15133	35
2000er	Nein	0,297	0,22171	12
	Ja	0,177	0,20131	7

N = 102

Interaktion zwischen Notierung und Time-Lag-Robustheit: Der Interaktionsterm einer gesättigten ANOVA ist nicht signifikant, wesentliche Zusammenhänge zwischen Notierung und Time-Lag-Robustheit sind nicht zu beobachten.

Interaktion zwischen Leistungsoperationalisierung und Time-Lag-Robustheit: Zu beachten ist hier einmal mehr die Verknüpfung zwischen der Leistungsoperationalisierung Experten und der Variablen Time-Lag-Robustheit. Da annahmegemäß expertenbasierte Leistungsoperationalisierungen immer für Time-Lags kontrollieren wenn nicht explizit nach vergangenheitsbezogenen Daten gefragt wird, ist der Zusammenhang zwischen Time-Lag-Robustheit und der Leistungsoperationalisierung Experten nicht sinnvoll interpretierbar. Entsprechend reduziere

ich den Datensatz durch Ausschluss aller Fälle, die mit expertenbasierten Leistungsoperationalisierungen arbeiten. Die nach der Eliminierung verbleibende Interaktionsmatrix ist im Folgenden dargestellt.

Tabelle 37: Interaktionsmatrix Leistungsoperationalisierung / Time-Lag Robustheit

Leistungsoperationalisierung	Robust für Zeitversatzproblem	Mittelwert	Standardabweichung	N
Accountingbasiert	Nein	0,160	0,17148	24
	Ja	0,176	0,14634	23
Marktbasiert	Nein	0,052	0,15960	10
	Ja	0,192	0,17039	11
Prozessbasiert	Nein	0,257	0,20888	22
	Ja	0,252	0,15564	13

Bei allen Leistungsoperationalisierungen zeigen Time-Lag robuste Kennzahlen höhere Korrelationen an als nicht robuste, mit Ausnahme der prozessbasierten, hier gibt es keinen nennenswerten Unterschied. Der ANOVA-Interaktionsterm ist nicht statistisch signifikant. Allerdings ist festzuhalten, dass der Beitrag der Time-Lag Effekte zur Varianzaufklärung der abhängigen Variablen „Korrelation zwischen Performance und Ressource“ massiv sinkt, wenn expertenbasierte Operationalisierungen aus den ANOVA-Modellen entfernt werden. Auch hier zeigt sich also, dass der Einfluss von Time-Lags im Vergleich zu dem der Leistungsoperationalisierung gering bis marginal ausfällt.

Interaktionseffekte höherer Ordnung: Die Berechnung von Interaktionseffekten höherer Ordnung ist aufgrund der begrenzten Größe des Datensatzes und daraus resultierenden beschränkten Gruppengrößen nicht mehr sinnvoll möglich.

5.5 Diskussion der Ergebnisse

Tabelle 38: Überblick Hypothesentests Metaanalyse

Nr.	Hypothese	Ergebnis
H1.1	<i>Korrelationen zwischen Leistung und Ressourcen unterscheiden sich signifikant, je nachdem, welche Branchen betrachtet werden.</i>	nicht bestätigt
H1.2	<i>Studien, die mehr als eine Branche umfassen, berichten geringere Korrelationen und weisen höhere Varianzen auf als Studien, die sich auf eine oder wenige Branchen konzentrieren.</i>	nicht bestätigt
H1.3	<i>Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, welche Zeitabschnitte in den zu Grunde liegenden Primärstudien betrachtet werden.</i>	Keine Aussage möglich, von H2 überlagert

H1.4	<i>Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, ob gelistete oder nicht gelistete Unternehmen untersucht werden.</i>	Keine Aussage möglich, von H2 überlagert
H2	<i>Die Korrelation zwischen Performance und Ressource fällt schwächer aus, je weiter sich die zu Grunde gelegte Performanceoperationalisierung vom Konzept der Kernperformance entfernt, je mehr sie also bereits Objekt des Verteilungskampfes zwischen Stakeholdergruppen war.</i>	bestätigt
H3.1	<i>Studien, die mit Time-Lag robusten Leistungsoperationalisierungen arbeiten, finden einen stärkeren Zusammenhang zwischen Ressourcen und Performance als Studien, deren Performanceoperationalisierungen nicht robust gegenüber Zeitversatzphänomenen sind.</i>	Teilweise bestätigt
H3.2	<i>Der Teil der Studien, der nicht mit für Zeitversatz robusten Verfahren arbeitet, weist eine höhere Varianz in der abhängigen Variablen auf als der Teil, der robuste Verfahren einsetzt.</i>	nicht bestätigt

5.5.1 Populationsselektion

Branche

Ein signifikanter Einfluss der Branche auf Studienergebnisse konnte nicht festgestellt werden. Weder das technologische Regime noch die Unterscheidung, ob produzierende Unternehmen oder Dienstleister untersucht wurden, hat einen nachweisbaren Einfluss auf die Stärke des Zusammenhangs zwischen Ressourcen und Performance. Tests auf der Ebene einzelner Branchen waren mit dem vorliegenden Datensatz nicht möglich. Im Sample liegt ein Datensatz (in zwei Artikeln) vor, der zur Erklärung dieses Ergebnisses nützlich ist (Powell, 1992a; Powell, 1992b). In beiden Fällen beleuchtet Powell den Einfluss koordinierten strategischen Denkens des Managements, einer Ressource im Sinne des Resource-based view, auf die Leistung von Organisationen, gemessen als Einschätzung interner Experten. Das Sample besteht aus zwei Branchen, Frauenbekleidung und Holzmöbel. In einem Fall wird aus den beiden Branchen ein kumuliertes Sample gebildet, die Korrelation zur Performance liegt hier bei 0,20, die zweite Studie disaggregiert nach Branchen. In der Frauenbekleidungsindustrie liegt die Korrelation zwischen sorgfältiger Strategieausarbeitung und Performance bei 0,06, in der Holzmöbelindustrie bei 0,34. Im arithmetischen Mittel ergibt sich wieder die Korrelation zur Performance von 0,20. Der Unterschied zwischen beiden Branchen liegt in der Umweltdynamik. Während die Holzmöbelindustrie in dieser Beziehung sehr stabil ist, strategisches Arbeiten sich also über die Zeit auszahlen kann, ist die Frauenbekleidungsindustrie von einer sehr hohen Dynamik gekennzeichnet, die langfristige Strategien kaum ergebniswirksam werden lässt (Powell, 1992b). Dieser Einzelfall liefert damit auch gleich eine mögliche Erklärung für die Nichtsignifikanz der Brancheneffekte in der vorliegenden Metaanalyse: Beide Branchen würden hier in dieselben Kategorien eingeordnet werden (Produktion beziehungsweise Low-Tech). Es lässt sich also annehmen, dass die Einteilung der Branchen in Produktion und Dienstleistungen zu grobkörnig war, um tatsächlich Brancheneffekte aufspüren zu können. Gleichzeitig bietet der

Datensatz der Metaanalyse keine Möglichkeit einer feinkörnigeren Branchenbetrachtung, die Frage nach der Stärke des Brancheneinflusses muss also unbeantwortet bleiben.

Überraschen muss jedoch, dass das technologische Regime keinen messbaren Einfluss auf den Zusammenhang zwischen Ressourcen und Performance ausübt. Die empirische Evidenz in der Literatur für Unterschiede zwischen Low- und High-Tech ist überzeugend (vergleiche beispielsweise Aggarwal & Hsu, 2009; Ang, 2008; Makri et al., 2006; Teece, 2000). Offenbar schlägt sich der Unterschied zwischen technologisch eher komplexen und eher weniger komplexen Organisationen bei höher aggregierten Zusammenhängen wie Ressourcen im Sinne des Resource-Based view und Leistung nicht mehr merklich nieder. Die Unterschiede scheinen vielmehr im Prozess der Entstehung von Ressourcen zu liegen und darin, *welche* Ressourcen in bestimmten technologischen Regimen zum Erfolg führen, nicht im grundsätzlichen Zusammenhang zwischen Ressourcen und Performance.

Grundsätzlich ist die Bewertung des Einflusses der Branche auf Zusammenhänge in der Strategischen Managementforschung aufgrund der Vielzahl der Faktoren, die durch die Branche in der engeren Umwelt einer Organisation beeinflusst werden, relativ komplex. Notwendig wäre eine tiefergehende Analyse der Branchenunterschiede, die mit dem vorliegenden Datensatz so nicht zu leisten ist. Datta, Guthrie & Wright (2005) beispielsweise unterteilen in ihrer Untersuchung zum Einfluss der Branche auf die Wirkung von High-Performance Worksystems die untersuchten Branchen anhand der FuE-Intensität, des Wachstums der Branche, des Produktdifferenzierungsgrades sowie der Branchendynamik (Zu- und Austritte aus dem Markt). Entsprechend komplexe Modelle sind zweifelsfrei notwendig, wenn kausale Aussagen zum moderierenden Einfluss der Branche auf Zusammenhänge in der Strategischen Managementforschung abgeleitet werden sollen. Als Ergebnis der Metaanalyse ist festzuhalten, dass der Zusammenhang zwischen Ressourcen und Performance nicht messbar von der hier vorgenommenen groben Brancheneinteilung abhängt.

Zeitraum

Die Hypothese, dass sich Zusammenhänge über die Zeit ändern können, lässt sich nicht eindeutig bestätigen. Zunächst wurden für die 1980er-Jahre deutlich geringere Korrelationen zwischen Ressourcen und Performance gemessen als dies in neueren Studien der Fall ist, auch feinkörnigere Einteilungen der Zeiträume, Konjunkturzyklen von etwa 5 Jahren Länge, lassen diese Unterschiede klar hervortreten. Allerdings lässt sich dieser Effekt auf die in den entsprechenden Zeiträumen wesentlich verwendeten Methoden zur Leistungsoperationalisierung zurückführen, wie die Analyse eines gesättigten Modells mit Zeitraum und Leistungsoperationalisierung zeigte. Dies bedeutet im Umkehrschluss auch keine Falsifikation der geprüften Hypothese, notwendig wäre die Betrachtung des Einflusses des Zeitraums auf einer ceteris paribus-Basis. Das heißt es müsste überprüft werden inwiefern sich die Korrelationskoeffizienten zwischen Jahrzehnten bei Studien unterscheiden, die mit derselben Leistungsoperationalisierung arbeiten. Dies lässt die Größe der resultierenden Cluster in den 1980er und 2000er-Jahren nicht zu. Explorative Indikationen lassen sich finden, diese sind aufgrund der dünn besetzten Cluster allerdings nicht als überzufällig einzuschätzen.

Damit ist die von Gadamer (1990) diskutierte Historizität der Erkenntnisse in der empirischen Sozialforschung (vergleiche Abschnitt 3.2.4) in der hier vorgelegten Studie nicht zu belegen. Zwar wäre bei Berechnung einer ANOVA mit lediglich einer Gruppierungsvariablen Zeitraum die Annahme noch naheliegend, dass die Unternehmen in den letzten Jahren ihre Ressourcen leistungsfördernder eingesetzt haben könnten als dies noch vor zwei Jahrzehnten der Fall war, möglicherweise sind auch Umweltveränderungen eingetreten, die heute eine andere Beziehung zwischen Ressourcen und Leistung bedingen als in vorhergegangenen Zeitabschnitten. Hierfür spräche die Beobachtung, dass sich die Zusammenhänge insbesondere im Zeitraum um 1990 herum stark anders darstellen als in den Vergleichszeiträumen, wie die Untersuchung auf Basis kürzerer Zeitfenster gezeigt hat. Die Turbulenzen um den Zerfall der Sowjetunion und das Ende des Ost-West-Konflikts könnten als Erklärung herangezogen werden warum in diesem Zeitraum andere Kausalitäten zum Zusammenhang zwischen Ressourcen und Performance galten. Dieser Erklärungsansatz wird jedoch durch das Verschwinden der Signifikanz der Gruppierungsvariable Zeitraum im gesättigten Modell in Frage gestellt. Damit ist keine Aussage möglich, ob ein solcher Effekt tatsächlich vorliegt, in der hier präsentierten statistischen Auswertung dominiert jedenfalls die Tatsache, dass in früheren Zeitfenstern Leistung häufiger marktbasierend operationalisiert wurde. Dies überstrahlt einen möglichen Effekt verschiedener Zeitfenster. Die Leistungsoperationalisierung erklärt klar mehr Heterogenität der Ergebnisse im Fachgebiet als der Zeitraum, in dem eine Primärstudie entstanden ist.

Als ein Indiz für eine mögliche Historizität von Forschungsergebnissen und Kausalitäten kann die Analyse der Interaktionsterme zwischen den Variablen Zeitraum und Notierung beziehungsweise Time-Lag-Robustheit verstanden werden. In den 2000er-Jahren kehren sich Muster, die in den 1980er und 1990er-Jahren noch zu beobachten waren, um, insbesondere sind die Beziehungen zwischen Ressourcen und Performance nun bei börsennotierten Unternehmen stärker ausgeprägt als bei nicht notierten. Dies kann auf eine tatsächliche Veränderung der Umweltbedingungen in den 2000er-Jahren hindeuten.¹²⁹ Ein Indiz für die Bedeutung des Zeitraums gibt auch die Studie von Miller & Shamsie (1996)¹³⁰. Die Autoren untersuchen den Einfluss der Bindung von Schlüsselpersonen (der Ressource) auf die Umsatzrentabilität (Leistungsoperationalisierung) von Filmstudios in zwei Zeitfenstern: Von 1936-1950 und von 1951-1965. Während im ersten Zeitfenster die Korrelation stark ausgeprägt ist (0,32) verschwindet sie im zweiten (0,01). Erklärt wird dies durch das stabile, beziehungsweise dynamische Umfeld, in dem die Gesellschaften in den entsprechenden Zeitfenstern jeweils tätig waren. Es liegt also zumindest anekdotenhafte Evidenz für die Relevanz von Zeiteffekten vor (vergleiche auch Abschnitt 3.2.4), auch wenn sich die Frage nach der Ursache der Veränderung der beobachteten Korrelationen aus der vorliegenden Analyse heraus nicht eindeutig beantworten lässt. Festzuhalten bleibt, dass sich die beobachteten Kausalitäten in unterschiedlichen Zeitfenstern in der Tat anders darstellten und so zumindest ein (geringer) Teil der Heterogenität der Ergebnisse im Forschungsfeld erklärt wird.

¹²⁹ Zu beachten sind jedoch auch hier die relativ geringen Fallzahlen für die 2000er-Jahre und die damit einhergehende relativ große Gefahr statistischer Fehlschlüsse.

¹³⁰ Die Studie wurde aufgrund der sehr geringen Fallzahlen aus dem Sample der Metaanalyse entfernt, vergleiche Abschnitt 5.4.1, und ist entsprechend auch eher als Fallstudie denn als quantitative Untersuchung zu verstehen.

Notierung

Hypothese H1.4 wurde im Rahmen der Metaanalyse zunächst bestätigt, die Beziehung zwischen Ressourcen und Performance unterscheidet sich zwischen börsennotierten und nicht börsennotierten Unternehmen. So finden Forscher bei nicht börsennotierten Organisationen deutlich stärkere Zusammenhänge zwischen Ressourcen im Sinne des Resource-based view und der Leistung dieser Organisationen. Allerdings lässt sich auch dieser Effekt auf unterschiedlich verwendete Leistungsoperationalisierungen zurückführen. Wie ausgeführt konzentrieren sich Samples, die börsennotierte Unternehmen erfassen, auf accounting- und marktbasierende Kennzahlen, während bei Studien, die nicht börsennotierte Unternehmen beleuchten, insbesondere expertenbasierte Kennzahlen Verwendung finden. Die Richtung dieser Kausalität ist eindeutig, da der Effekt, dass accounting- und marktbasierende Kennzahlen geringere Korrelationen zu Ressourcen im Sinne des Resource-based view anzeigen, in allen Samples gilt. Ähnlich wie im Fall des Zeitfensters dominiert auch hier der Effekt der Performanceoperationalisierung etwaige originäre Effekte der Sample-Selektion durch die Beschränkung der Grundgesamtheit auf börsennotierte Organisationen.

Auch dies bedeutet nicht, dass die Notierung keine Rolle spielt, generell ist auch hier zu beobachten, dass Studien, die mit nicht börsennotierten Unternehmen arbeiten, höhere Korrelationen zwischen Leistung und Ressourcen finden als solche, die sich ausschließlich auf börsennotierte Unternehmen konzentrieren. Zumindest bei prozess- und expertenbasierten Leistungsoperationalisierungen ist dies sichtbar, allerdings ist auch hier die Fallzahl zu klein um statistisch belastbare Aussagen über die mögliche Überzufälligkeit dieses Zusammenhangs treffen zu können.

Verschiedene Gründe sind denkbar, warum die Börsennotierung eine Rolle spielen könnte. Eine Erklärung könnte etwa sein, dass börsennotierte Organisationen im Mittel deutlich größer sind als nicht börsennotierte. Die Leistung großer Organisationen lässt sich schwerer einzelnen Faktoren wie einer bestimmten Ressource zuschreiben als dies bei kleineren Organisationen der Fall ist (vergleiche zu den Spezifika von Größe unter anderem Williamson (1999)).

Würdigung des Populationsselektionsproblems

Zusammenfassend lässt sich festhalten, dass die Sorge, durch eine Entkontextualisierung von Forschungsergebnissen systematisch Fehler zu generieren, berechtigt ist, der Effekt jedoch deutlich weniger stark ausgeprägt ist als die Frage, wie die Leistung einer Organisation gemessen wird. Zwei der drei getesteten Selektionskriterien sind statistisch signifikant und erklären damit einen Teil der Ergebnisvarianz in der gesuchten Kausalbeziehung, sie sind jedoch gleichzeitig auch nicht unabhängig von der Gruppierungsvariable Leistungsmessung und in ihrer Wirkung deutlich schwächer ausgeprägt. Die Frage, ob Ergebnisse in der Strategischen Managementforschung also tatsächlich nicht ohne weiteres über die Grenzen von Teilpopulationen hinweg übertragbar sind, lässt sich mit dem vorliegenden Datensatz nicht eindeutig beantworten, eine explorative Betrachtung legt dies aber nahe.

5.5.2 Appropriation

Hypothesentest

Klar ist zu erkennen, dass die Leistungsoperationalisierung als Einflussgröße in der Tat einen bedeutenden Teil der Varianz in den Kausalbeziehungen zwischen Performance und Ressourcen erklären kann, deutlich mehr als die anderen beiden diskutierten Methodenprobleme, Populationsselektion und Time Lags. Die Art des Einflusses deckt sich dabei mit dem Bild vom Prozess der Wertschöpfung und Wertverteilung in Organisationen, das die Appropriationstheorie zeichnet. Wie in der konzeptionellen Betrachtung entwickelt (vergleiche Abschnitt 3.3) und in Hypothese H2 ausformuliert lassen sich nicht nur Unterschiede zwischen appropriationsrobusten und nicht robusten Kennzahlen ausmachen, wie Crook et al. (2008) bereits zeigen konnten, sondern es ist darüber hinaus zu beobachten, dass Performanceoperationalisierungen, die näher am entwickelten Konzept der Kernperformance (vergleiche Abschnitt 3.3.4) liegen, klar überzufällig stärkere Zusammenhänge zwischen Ressourcen und Leistung abbilden als solche, die erst nach dem Wertverteilungsprozess in Organisationen messen.

Gruppiert man das Sample der Metaanalyse nach Indikatorengruppen ist die Situation eindeutig: Marktbasierte Operationalisierungen zeigen die niedrigsten Korrelationen zwischen Leistung und Ressourcen an, gefolgt von accountingbasierten, prozessbasierten und schließlich expertenbasierten oder subjektiven Kennzahlen. Die Analyse der Interaktionsbeziehungen zu den anderen untersuchten Gruppenvariablen zeigt, dass dieser Effekt auch nicht durch Einflüsse anderer in die Metaanalyse einbezogener Variablen erklärbar ist.

Auf der Ebene einzelner Leistungsoperationalisierungen fällt negativ auf, dass der vorliegende Datensatz, obwohl relativ umfangreich, weniger Bandbreite an Leistungsoperationalisierungen abbildet als dies nach den Ergebnissen der qualitativen Literaturreview zu erhoffen gewesen wäre. Leistungsoperationalisierungen sind also offenbar deutlich mit bestimmten Forschungsfeldern assoziiert, Studien auf Basis des Resource-based view arbeiten im Vergleich zum in Kapitel 4 entwickelten Gesamtbild deutlich häufiger mit an Key Informants adressierten Fragebogenerhebungen, dafür sind marktbasierte Kennzahlen die Ausnahme. Andere Operationalisierungen wie Überleben/Scheitern, in der Literaturreview die fünfthäufigste Operationalisierung, finden im Rahmen der Untersuchungen zum Resource-based view nahezu keine Anwendung. Der Anteil der accountingbasierten Leistungsoperationalisierungen entspricht mit einem Drittel in etwa dem Niveau, das auch in der Literaturreview zu beobachten war. Auf der Ebene der Einzeloperationalisierungen lassen sich mit dem vorliegenden Datensatz lediglich sechs Leistungsindikatoren sinnvoll im Rahmen der Metaanalyse auswerten. Diese Einzelauswertung zeigt das Bild, das im Rahmen der konzeptionellen Betrachtungen zur Appropriationstheorie vorhergesagt wurde und das sich bereits bei der Analyse der aggregierten Operationalisierungen abgezeichnet hat: Tobins q, die einzige marktbasierte Größe mit befriedigender Fallzahl, misst die geringsten Korrelationen zwischen Ressourcen und Performance, gefolgt von den beiden accountingbasierten Leistungsoperationalisierungen Gewinn und Kapitalrentabilität. Diese drei Größen liegen mit 0,09 bis 0,18 deutlich unter dem Gesamtmittel für die abhängige Variable von 0,23. Die beiden prozessbasierte Größen, Produktivität und Umsatz, messen Werte über dem Mittel für das Gesamtsample (0,26 und 0,27), während die Leistungserfassung durch die Befragung von organisationsinternen Experten deutlich höhere Werte erzeugt (0,35). Auffällig ist in dieser Einzelbetrachtung, wie homo-

gen die Gruppen hier jeweils sind, eine starke empirische Unterstützung der Appropriationstheorie.

Im Kontext der Problemstellung dieser Arbeit ist also festzuhalten, dass das Ergebnis beispielsweise auch der Metaanalysen zum Resource-based view und damit die hieraus abgeleiteten Aussagen mit all ihren Implikationen für Praxis und weitere Forschung massiv von der zu Grunde gelegten Performanceoperationalisierung abhängt. Würde man die Untersuchung zum Zusammenhang zwischen Performance und Ressourcen ausschließlich auf Basis von markt- und accountingbasierten Kennzahlen durchführen wäre, mit einem bereits für methodische Artefakte (vergleiche Hunter & Schmidt, 2007) nach oben korrigierten Mittelwert in der Korrelation von unter 0,15, kaum eine Annahme der Hypothese des Resource-based view, nämlich die Konstatierung eines signifikanten Zusammenhangs zwischen Ressourcen und Performance, möglich. Ohne Einbeziehung der statistischen Korrekturfaktoren für Messfehler läge die Korrelation lediglich bei 0,11. Prozessbasierte Kennzahlen dagegen zeigen deutlich höhere Korrelationen, hier wäre eine Annahme der Hypothese, dass Ressourcen im Sinne des Resource-based view ceteris paribus zu höherer Leistung der entsprechenden Organisation führen, vertretbar. Aus Sicht der in dieser Arbeit leitenden Forschungsfrage ist nun festzustellen, dass es sich bei der relativen Verhandlungsmacht der Stakeholdergruppen also tatsächlich um eine unbeobachtete Moderatorvariable handelt, die die Heterogenität der Forschungsergebnisse in der Strategischen Managementforschung in Teilen erklären kann. Im Rahmen der hier vorgelegten Metaanalyse konnten auf der disaggregierten Ebene nur Performanceoperationalisierungen untersucht werden, die die Appropriationsleistung der Anteilseigner erfassen, keine, die spezifisch die Rentenappropriation einer anderen Stakeholdergruppe, etwa des Managements oder der Arbeitnehmer, beleuchten. Bei diesen wären ähnlich niedrige Ausprägungen der abhängigen Variablen zu erwarten. Diese Vermutung zu überprüfen muss zukünftigen Forschungen vorbehalten bleiben. Die Appropriationsleistung anderer Stakeholder als der Anteilseigner zu messen ist, wie in Abschnitt 3.3 ausführlich diskutiert, komplex.

Betrachtet man das Ergebnis der vorliegenden Metaanalyse in Bezug auf den Ressourcenbasierten Ansatz kann festgehalten werden, dass Ressourcen also tendenziell zwar die Kernwertschöpfung der Organisation erhöhen, von diesen zusätzlich geschaffenen Werten allerdings nur wenig bei den Anteilseignern ankommt. Andere Stakeholdergruppen sind unter Umständen deutlich erfolgreicher darin, sich Werte anzueignen. Dies bedeutet auch, dass Ressourcen möglicherweise dazu geeignet sein können, die relative Verhandlungsmacht der Stakeholdergruppen zu verschieben. Hier sind weitere Forschungen notwendig, um diese Effekte genauer und insbesondere spezifisch für einzelne Ressourcen zu beleuchten. Klar wird auch, dass die bisherigen Ansätze zur metaanalytischen Überprüfung des Resource-based view (Newbert, 2007, Crook et al., 2008) in ihren gegenläufigen Einschätzungen zur empirischen Relevanz des Resource-based view ein zu wenig differenziertes Verständnis der Heterogenität der betrachteten Primärstudien und des Konzepts „Leistung“ zu Grunde legten. Die für die hier vorgelegte Metaanalyse verständnisleitende Appropriationstheorie ermöglicht ein besseres Verständnis des Zusammenhangs zwischen Ressourcen im Sinne des Resource-based view und verschiedenen Dimensionen der Leistung einer Organisation. Ansätze vergleichbarer Überlegungen finden sich bereits in Crook et al. (2008), mit der hier vorgelegten Metaanalyse liegt jedoch nun zum ersten Mal empirische Evidenz für den Einfluss des Wert-

verteilungsprozesses, wie ihn die Appropriationstheorie beschreibt, auf den Zusammenhang zwischen Ressourcen und Performance vor.

Die Leistungsmessung durch „Experten“

Etwas überraschen muss das Ausmaß, in dem sich die Operationalisierung Experten von den „harten“, in der Terminologie von Rowe & Morrow Jr. (1999) „objektiven“, Leistungsoperationalisierungen abhebt. Insbesondere die in Abschnitt 3.3.3.2 herausgearbeitete Vermutung relativ starker Korrelationen von Experteneinschätzungen mit accountingbasierten Kennzahlen ist eindeutig zu verwerfen. Die Einschätzung von Experten zur Leistung von Organisationen korreliert, wie gezeigt wurde, sehr viel stärker mit dem Vorhandensein von Ressourcen im Sinne des Resource-based view als andere Leistungsoperationalisierungen. Wie ist dies zu erklären?

Grundsätzlich ist aus Sicht der Appropriationstheorie durchaus zu erwarten, dass Experten in der Lage sein könnten, eine generalistische Perspektive einzunehmen. Sie sollten aufgrund ihrer privilegierten Positionierung innerhalb der Organisation¹³¹ und des damit verbundenen besseren Zugangs zu Informationen (vergleiche Steigenberger, Pawlowsky, Lübcke & Schulte, 2010; Wall et al., 2004) einen relativ guten Überblick über die Gesamtleistung einer Organisation haben und die Auswirkungen von Ressourcen auf die Leistung gut und von moderierenden Einflüssen unverfälscht abschätzen können. Aus einer sozialkonstruktivistischen Sicht sind Experten gut in der Lage, den Gesamterfolg realistisch einzuschätzen, da sie die Wirklichkeitskonstruktionen der Organisation, der sie angehören, teilen. Nach Lebas & Euske (2007) ist es ihnen so möglich, die Handlungen von Akteuren in und um ihre Organisation abzuschätzen und sich so ein wirklichkeitsnahes, das heißt in diesem Fall vom politischen Prozess der Appropriation unverzerrtes Bild des Erfolgs eines Unternehmens zu machen. Diese Effekte scheinen stärker zu wiegen als die geistige Nähe der regelmäßig befragten Gruppe des Top-Managements zu accountingbasierten Kenngrößen, die noch in Abschnitt 3.3.3.2 vermutet wurde.

Allerdings werden gegen diese Argumentation, die den Experten eine gute Einschätzung der Leistung von Organisationen bescheinigt, nachdrücklich Gegenargumente vorgebracht. Zunächst ist hier das Problem zu nennen, dass Experten häufig retrospektiv, also zu Ereignissen in der Vergangenheit, befragt werden (March & Sutton, 1997). Dies führt zu Unzuverlässigkeit in den Aussagen, da die Vergangenheit subjektiv immer von der Gegenwart aus konstruiert wird und über Experten so nicht wirklichkeitsgetreu wiedergegeben werden kann. Dies könnte als Verzerrung die Möglichkeit der Etablierung starker Korrelationen behindern (vergleiche

¹³¹ Die folgenden Ausführungen beziehen sich, wenn nicht anders angegeben, auf organisationsinterne Experten. Grundsätzlich können Experten natürlich auch außerhalb der Organisation stehen, diese Art der Leistungsmessung spielt in der Strategischen Managementforschung jedoch, wie die Literaturreview in Abschnitt 4.2.2 und auch die hier vorgelegte Meta-Analyse zeigen, kaum eine Rolle. Grundsätzlich gelten die hier präsentierten Überlegungen in etwas abgeschwächter Form auch für organisationsexterne Experten. Externe Experten sollten etwas schlechter in Informationsströme eingebunden sein, dafür ist das unten angesprochene Problem der Eigeninteressen weniger relevant. Vergleiche zur Rolle von Experteneinschätzungen und zur Arbeit mit Experten in der Leistungsmessung insbesondere auch Steigenberger, Pawlowsky, Lübcke & Schulte, 2010.

auch Starbuck, 2004). Tversky & Kahneman (1973) sprechen in diesem Kontext vom availability bias, Personen tendieren dazu, bestimmte Umstände, insbesondere solche, die erst kurz zurückliegen oder häufig auftreten, wesentlich leichter zu erinnern als andere, länger zurückliegende oder selten auftretende. Shanteau (1992) dagegen führen als besondere Qualität eines Experten an, dass sein Wissen und seine Erfahrung ihm eben diese Erfassung breiter Zusammenhänge ohne wesentlichen availability bias ermöglichen. Diese Frage tiefer zu elaborieren liegt jenseits des Fokus dieser Arbeit, vergleiche zur Entscheidungsfindung von Experten weiterführend beispielsweise Klein (1995), Mosier (2008), den vor kurzem erschienenen Herausgeberband von Schraagen, Militello, Ormerod & Lipshitz (2008) und die Arbeiten von Gerd Gigerenzer, etwa Gigerenzer & Goldstein (1996).

Es bleibt festzuhalten, dass sich bislang keine einheitliche Meinung zur grundsätzlichen Befähigung von Experten, befriedigend objektive Bewertungen vorzunehmen, herausgebildet hat; auch die empirische Basis ist uneindeutig, zumal der Maßstab dafür, welches Ausmaß an Objektivität als befriedigend anzusehen ist, je nach Autor und Fragestellung schwankt. Die Annahme, dass Experten in der Lage sind Entscheidungs- und Bewertungsheuristiken einzusetzen, die denjenigen von Nichtexperten deutlich überlegen sind, kann jedoch als stark fundiert angesehen werden.¹³² Entsprechend scheint ein gewisses grundsätzliches Zutrauen in Expertenurteile zur Bewertung von Leistungen derjenigen Organisationen, denen sie angehören, gerechtfertigt. Voraussetzung hierfür ist natürlich, dass tatsächlich *Experten* die entsprechenden Angaben zur Leistung machen, dies wird regelmäßig im in der Strategischen Managementforschung üblichen postalischen Fragebogendesign nicht zu überprüfen sein.¹³³ Die Tatsache, dass in der Metaanalyse mit Expertenurteilen hohe Korrelationskoeffizienten zwischen Ressource und Leistung verbunden sind, deutet darauf hin, dass Experten eher weniger Moderatorvariablen ausgesetzt sind als dies im Fall anderer Leistungsoperationalisierungen der Fall ist. Ein availability bias, der zu schwächeren oder zumindest inkonsistenten, das heißt mit hohen Varianzen einhergehenden, Korrelationen führen müsste, ist empirisch nicht zu beobachten.

Der zweite Einwand, den insbesondere March & Sutton (1997) gegen die Nutzung von Experten in der Strategischen Managementforschung vorbringen zielt auf die persönliche Motivationslage der (organisationsinternen) Experten ab. Als Mitglieder der Organisation, die es zu bewerten gilt, sind sie nicht unvoreingenommen in der Evaluation der Leistung dieser Organisation. Vielmehr werden sie dazu tendieren, Informationen zu generieren, die ein positives Selbstbild stützen. Sie werden die Leistung ihrer Organisation also tendenziell gut, annahmegemäß zu gut einschätzen. Letztere Vermutung ist nicht von der Hand zu weisen, für den hier untersuchten Zusammenhang auf der Basis von Korrelationen ist eine Über- oder Unterschätzung einer gegebenen objektiven Wahrheit jedoch nicht notwendigerweise schädlich. Eine hohe Korrelation ergibt sich nicht dann, wenn eine möglichst hohe Ausprägung der abhängigen Variablen Leistung gemessen wird sondern dann, wenn über die Fälle des Samples hohe Ausprägungen der abhängigen Variablen mit eben solchen der erklärenden Variablen, der jeweiligen Ressource, korrespondieren und umgekehrt. Eine potentiell zu gute Einschätzung

¹³² Zur Abgrenzung von Experten zu Nichtexperten vergleiche Steigenberger, Pawlowsky, Lübcke & Schulte, 2010.

¹³³ Vergleiche zum Vorgehen bei Fragebogenerhebungen in der Strategischen Managementforschung beispielsweise auch Slater & Athuahene-Gima, 2004.

der Leistung durch Experten könnte im Sinne der Argumentation dieser Arbeit die Korrelationen also nicht verzerren, sofern das Ausmaß der Überschätzung der eigenen Leistung über die einzelnen Fälle, beziehungsweise die einzelnen Experten im Sample nur zufällig und nicht systematisch schwankt.¹³⁴ Veranschaulichend lässt sich die Argumentation der Appropriationstheorie ja auch so umformulieren: Je mehr unerklärte Faktoren, etwa die Verhandlungsmacht einer bestimmten Stakeholdergruppe, in die Operationalisierung der abhängigen Variablen einbezogen sind, desto geringer die beobachtete Korrelation mit der erklärenden Variablen.

Vor diesem Hintergrund sind die hohen Korrelationen im Rahmen der Nutzung von Experteneinschätzungen erklärbar. Wenn Lebas & Euske (2007) recht haben, müsste genau dieser Fall eintreten: Die Experten können die Leistung von Organisationen abstrahiert von Verteilungskämpfen einschätzen, da sie die Akteure und Akteursgruppen in ihren Organisationen kennen und deren Handlungen, die lediglich der Verteilung der Kernperformance dienen, aus ihrer Einschätzung der Gesamtleistung ihrer Organisation eliminieren können. Diese Bewertung trägt auch wenn anstatt des sozialkonstruktivistischen Leistungsverständnisses von Lebas & Euske (2007) das dieser Arbeit zu Grunde liegende kritisch-rationalistische Verständnis erkenntnisleitend ist. Über diese Fähigkeit zur „Reinigung“ der Leistung von leistungsfremden Einflüssen wie Verteilungskämpfen verfügt keine andere Performanceoperationalisierung in diesem Maße. Damit wird das hohe Niveau der Korrelationskoeffizienten verständlich, wenn die Grundannahme gilt, dass es tatsächlich (im kritisch-rationalistischen Sinn) eine starke Korrelation zwischen Leistung und Ressourcen gibt.

Ein einfaches fiktives Rechenbeispiel soll diese Argumentation illustrieren. Gegeben seien drei börsennotierte Unternehmen A, B und C, die in unterschiedlichen Branchen tätig sind. Diese Unternehmen verfügen über die interessierende Ressource in den Ausprägungen 4, 6 und 8. Unternehmen A ist von seinen Abnehmern stark abhängig, Unternehmen B ist von einer starken Arbeitnehmerschaft geprägt während Unternehmen C durch Turbulenzen bei einem Konkurrenzunternehmen an der Börse deutlich abgewertet wurde. Annahmegemäß besteht ein starker wahrer und der Einfachheit halber linearer Zusammenhang zwischen der interessierenden Ressource und der Kernperformance, eine hohe Ausprägung der Ressource führt also ceteris paribus zu hoher Leistung und umgekehrt, die wahre Korrelation beträgt also 1. Die Leistung wird nun mit der marktbasierter Kennzahl TSR, dem accountingbasierten ROI und der prozessbasierten Operationalisierung Umsatz gemessen, zudem wird das Top-Management als Experten nach seiner Einschätzung zur relativen Performance ihrer Organisation in Relation zu einer Branchen-Vergleichsgruppe befragt. Alle Leistungsoperationalisierungen sind der Einfachheit halber auf eine Skala von 1 bis 10 normiert. Folgende Tabelle gibt die angenommenen Messwerte wieder:

¹³⁴ Solche systematischen Schwankungen könnten durch die anderen beiden betrachteten Methodenprobleme verursacht werden, beispielsweise wäre es denkbar, dass Key Informants in stark in der Öffentlichkeit stehenden Branchen die Leistungen ihrer Organisation stärker überschätzen als solche in Branchen, in denen sich Manager weniger in der Öffentlichkeit rechtfertigen müssen. Belege für solche Verzerrungen lassen sich mit den Daten der Metaanalyse aber nicht finden.

Tabelle 39: Rechenbeispiel Appropriation

Unternehmen	Ressource	TSR	ROI	Umsatz	Experten	Kernperformance
A	4	2	3	3	6	4
B	6	3	3	6	8	6
C	8	2	6	8	10	8
Korrelation		-0,54	0,58	0,94	1	1

Die Marktmacht der Abnehmer führt bei Unternehmen A dazu, dass Umsatz und Gewinnmargen relativ gering ausfallen, dies wird an der Börse nicht goutiert; die starke Arbeitnehmerschaft bei B führt zu einer hohen Appropriationsleistung der Mitarbeiter und entsprechend geringem ROI und einer eher negativen Wahrnehmung an der Börse; bei Unternehmen C fällt aufgrund der oben beschriebenen Situation lediglich der TSR gering aus. Die Experten sind in der Lage, die individuellen Probleme zu erkennen, einzuordnen und die Leistung mit einem Wert auszugeben, der der Kernperformance grob entspricht. Allerdings überschätzen sie zur Aufrechterhaltung ihres positiven Selbstbildes die Leistung ihrer Organisationen deutlich.

Der Korrelationskoeffizient nach Pearson (vergleiche beispielsweise Brosius, 2008, 507), der für wenigstens intervallskalierte Daten Verwendung findet, berechnet sich folgendermaßen:

$$r = \frac{\frac{1}{N-1} * \sum_{i=1}^N (X_i - \bar{X}) * (Y_i - \bar{Y})}{\sqrt{\frac{1}{N-1} * \sum_{i=1}^N (X_i - \bar{X})^2} * \sqrt{\frac{1}{N-1} * \sum_{i=1}^N (Y_i - \bar{Y})^2}}$$

N ist hierbei die Zahl der Fälle, hier also 3, \bar{X} und \bar{Y} sind die Mittelwerte von Ressource beziehungsweise Leistungsoperationalisierung über die drei Fälle, also 6 für die Ressourcen beziehungsweise 2,3, 4, 5 und 8 für die vier Performanceoperationalisierungen; X_i und Y_i sind die einzelnen Ausprägungen von Ressource beziehungsweise Performance. Die tatsächliche Korrelation wäre, da X und Y annahmegemäß linear voneinander abhängen, 1. Das Ergebnis der Berechnungen ist der obigen Tabelle zu entnehmen.

In diesem Rechenbeispiel beträgt die Korrelation im Fall der Experten 1, da die Abstände zwischen den Ausprägungen der Variablen und ihren Mittelwerten bei beiden Variablen, den Ressourcen und der Performance, identisch sind. Die Experten wären hier also in der Lage gewesen, die tatsächliche Korrelation perfekt abzubilden. Gleichzeitig ist zu erkennen, dass die Höhe der Messwerte keine unmittelbaren Auswirkungen hat. Die Werte der Experten liegen durchgängig höher als die „wahren“ Werte (Spalte Kernperformance), da lediglich die Abstände zum Mittelwert berechnet werden ist dies nicht schädlich.

Wie das Beispiel zeigt deuten die beobachteten Einflüsse der Leistungsoperationalisierungen auf die Höhe der Korrelationskoeffizienten in der Metaanalyse auf das zunehmende Vorliegen unbeobachteter Einflüsse hin, je weiter sich die Leistungsoperationalisierung vom Konzept der Kernperformance entfernt. Geht man nun davon aus, dass im Fall des in dieser Metaanalyse untersuchten Feldes ein tatsächlicher Zusammenhang zwischen Ressourcen und Performance besteht, ist zu konstatieren, dass Experten diesen offenbar am besten abbilden, da sie den wenigsten unerklärten Einflussfaktoren unterliegen. Die Appropriationstheorie liefert einen überzeugenden Erklärungsansatz für diese Beobachtung.

Die einzige mit den vorliegenden konzeptionellen und empirischen Daten denkbare alternative Erklärung für die hohen Korrelationen wäre, dass die Experten selbst bereits eine Verbindung zwischen den in ihrer Organisation vorliegenden Ressourcen und der Leistung ihrer Organisation ziehen, AV und UV damit nicht unabhängig wären. Wäre dem so, läge ein logischer Zirkelschluss vor, Korrelationen zwischen Ressourcen und dieser Performanceoperationalisierung könnten dann nicht mehr sinnvoll berechnet werden, beziehungsweise wären die hier vorliegenden dann aussagegelos. Argumente oder Studien, die diesen Verdacht belegen könnten, sind mir jedoch nicht bekannt, auch die Konstruktion der Befragungen in den Primärstudien deutet nicht darauf hin, dass dieser Fall wahrscheinlich ist. Vielmehr werden die Experten in der Regel anhand einer Reihe von Items nach unterschiedlichen Dimensionen der Leistung der Organisation befragt (vergleiche die einzelnen Primärstudien, Tabelle A.2 im Anhang). Ganz kann der Verdacht auf tautologische Schlüsse im Fall der Leistungsoperationalisierung Experten aber nicht ausgeräumt werden. Die vorliegende Arbeit bietet jedoch keine methodischen Anhaltspunkte zur Unterstützung dieser Vermutung, weitere Forschungen sind hier dennoch notwendig.

5.5.3 Time-Lag

Die Metaanalyse zeichnet ein uneinheitliches Bild zum möglichen Einfluss des Time-Lag Problems (Abschnitt 5.4.3.3). Korrelationen zwischen Leistung und Ressourcen fallen schwächer aus, wenn Primärstudien mit nicht für Zeitversatzprobleme robusten Verfahren arbeiten. Studien, die mit für Zeitversatz robusten Leistungsoperationalisierungen arbeiten, finden höhere Korrelationskoeffizienten. Dieser Zusammenhang ist allerdings, wie die Berechnung der Interaktionsterme zeigt, schwach. Kontrolliert man für unterschiedliche Leistungsoperationalisierungen fällt auf, dass das Time-Lag Problem insbesondere bei marktbasieren Kennzahlen einen erheblichen Einfluss ausübt, bei accountingbasierten noch zu beobachten ist und bei prozessbasierten verschwindet. Für diese Beobachtung liegt keine klare konzeptionelle Erklärung vor, notwendig wäre auch hier eine Dekomposition der Kategorie „prozessbasiert“, dies wird eine Aufgabe für Folgeforschungen sein. Die aus den statistischen Auswertungen ableitbare Stärke des Zeitversatzeffekts ist in jedem Fall sehr gering. Entsprechend homogenisierte Teilsamples weisen nur unwesentlich geringere Varianzen auf als das Gesamtsample, der prognostizierte Zusammenhang gilt überdies nur in zwei der drei betrachteten Zeitfenster. Es ist also festzuhalten, dass die in Abschnitt 3.4 entwickelte Hypothese zur Relevanz des Time-Lag Problems nur teilweise bestätigt werden kann.

Betrachtet man in Ergänzung zur Literaturreview noch die Häufigkeiten, mit der mit für Zeitversatz robusten und nicht robusten Verfahren gearbeitet wurde, fällt auf, dass in der Metaanalyse, die Studien von 1991 bis 2006 erfasst, deutlich weniger Analysen mit timelagrobusten Verfahren durchgeführt wurden als dies im Sample der Literaturreview, das die Zeit von 2006 bis 2008 betrachtet, der Fall ist. Auch wenn die beiden Samples natürlich in Aufbau und Erkenntnisinteresse nicht identisch sind, unterstützt dies doch die im Rahmen der Literaturreview aufgeworfene Vermutung, dass in den letzten Jahren eine Sensibilisierung für die Zeitversatzproblematik stattgefunden hat. Wenn dies der Fall ist, ist diese Sensibilisierung aber erst in jüngster Zeit aufgetreten, betrachtet man die Interaktionsmatrix zwischen Time-Lag und Zeitraum im Rahmen der Metaanalyse ist für die erste Hälfte der 2000er-Jahre keine wesentliche Erhöhung des Anteils an Time-Lag-robusten Operationalisierungen im Vergleich

zu den Vorperioden zu verzeichnen. Über die verschiedenen Performanceoperationalisierungen ist, wie die Interaktionsmatrix zeigt, die Verwendung von robusten Operationalisierungen weitgehend konstant, Studien, die mit marktbasieren oder prozessbasierten Operationalisierungen arbeiten beziehen das Time-Lag Problem nicht wesentlich häufiger oder seltener mit ein als solche, die accountingbasierte Kennzahlen nutzen.

5.6 Abgrenzung und Limitationen

5.6.1 Limitationen

Die vorliegende Untersuchung weist einige Schwächen auf, auf die hinzuweisen geboten erscheint. Zunächst ist die Etablierung eines Samples ohne Anwendung des Vier-Augen-Prinzips immer bedenklich, die sogenannte Interrater Reliability, von Orwin (1994) als zentraler Qualitätsindikator für die Zuverlässigkeit von Kodierungen im Rahmen der Metaanalyse diskutiert, kann so nicht ermittelt werden. Auf diese Problematik und den gewählten Lösungsweg bin ich bereits in Abschnitt 5.4.1 ausführlich eingegangen, das Problem sollte durch die Verwendung von bereits bestehenden Samples wesentlich eingegrenzt sein.

Im Rahmen der Sampleauswahl ist zu beachten, dass lediglich Studien in die Metaanalyse einbezogen wurden, die vor 2008 erschienen sind. Fraglich ist, inwieweit damit die Annahme der Repräsentativität des Samples für die Grundgesamtheit der Studien, die den Zusammenhang zwischen Ressourcen und Performance untersuchen, verletzt ist. Repräsentativität ist eine Grundvoraussetzung zur Ableitung verallgemeinernder Aussagen aus einer Metaanalyse (Hedges, 1994) und überdies Bedingung für die Anwendbarkeit von fixed effects-Modellen, die in dieser Metaanalyse Verwendung fanden (vergleiche Hartung et al., 2009). Dies lässt sich aus dem vorliegenden Datenmaterial nicht letztgültig klären, doch scheint der subjektive Eindruck, den eine stichprobenhafte Überprüfung neuerer Arbeiten zum Resource-based view ergeben hat (vergleiche beispielsweise Becerra, 2008; Newbert, 2008; Pacheco-de-Almeida et al., 2008; Holcomb, Holmes Jr. & Connelly, 2009; Morgan et al., 2009), darauf hinzuweisen, dass grundlegende Diskussionen um den Bestand des Theoriegebäudes und um den Zusammenhang zwischen Performance und Ressourcen im Sinne des Resource-based view nicht mehr geführt werden. Aus diesem Grund scheint die Annahme, dass die wesentlichen Diskussionen um den Zusammenhang zwischen Performance und Ressourcen zwischen 1991, dem Zeitpunkt des Erscheinens von Barneys grundlegendem Artikel (Barney, 1991) und 2007, dem Ende des Betrachtungszeitraums der Metaanalyse, geführt wurden, berechtigt. Entsprechend gehe ich davon aus, dass die gewonnenen Ergebnisse sich bei Einbeziehung weiterer, neuerer Arbeiten nicht wesentlich verändern würden.

Ein nicht lösbares Problem im Rahmen der Sampleauswahl ist die Korrektur für publication bias. Es ist lange bekannt dass tendenziell eher statistisch signifikante Zusammenhänge publiziert werden, während statisch nicht signifikante Beziehungen nur selten Eingang in wissenschaftliche Journals finden (vergleiche beispielsweise Cooper & Hedges, 1994; Hunter & Schmidt, 2007). Damit überschätzen Metaanalysen, die lediglich auf publizierten Journal-Artikeln basieren, wie die vorliegende, systematisch die „wahren“ Zusammenhänge. Dies erscheint im konkreten Fall der vorliegenden Arbeit allerdings relativ unkritisch, da ja primär nicht der Zusammenhang zwischen Performance und Ressourcen interessiert, sondern das

Augenmerk auf Faktoren gelegt wird, die zur Heterogenität der Ergebnisse in Untersuchungen dieses Zusammenhangs führen. Hinsichtlich der getroffenen Aussagen zum Resource-based view in Abschnitt 5.5 ist dieser Bias aber zu beachten.

Eine Schwäche der Studie kann auch darin gesehen werden, dass die vorliegende Metaanalyse die aufgeworfene Fragestellung nur anhand einer spezifischen Forschungsfrage, den empirischen Untersuchungen zum Resource-based view, bewertet. In anderen Settings könnten andere Ergebnisse vorliegen. Diese Limitation entspricht der allgemeinen Begrenzung der Aussagekraft, die mit Metaanalysen zu erreichen ist, wie von Hall & Tickle-Degnen (1994) diskutiert. Die Studie müsste in anderen Diskursräumen der Strategischen Managementforschung wiederholt werden um die hier erzielten Erkenntnisse zu validieren. Die damit einhergehenden Probleme (insbesondere Samplegröße) wurden in Abschnitt 5.3.1 bereits diskutiert.

Ein kleineres methodisches Problem stellt auch die Korrektur für unsystematische Messfehler durch einen Pauschalwert dar. Das Problem liegt insbesondere darin, dass relativ wenige Primärstudien Reliabilitätsmaße angeben und sich Reliabilitätsmaße mathematisch naturgemäß nur berechnen lassen, wenn ein Konstrukt mit mehreren Items gemessen wird. Die meisten Performanceoperationalisierungen, insbesondere accounting- und marktbasierter Kennzahlen, verschließen sich so der Reliabilitätsanalyse, da sie ja auf einem ganz konkreten Wert, etwa dem ROI oder dem TSR, basieren. Im konkreten Fall der vorliegenden Metaanalyse ist dieses Problem jedoch nicht allzu gravierend, da ja nicht die Korrelationskoeffizienten in ihrer Höhe interessieren, sondern Moderatorvariablen, die Unterschiede zwischen diesen Koeffizienten erklären sollen. Durch die Korrektur für unsystematische Messfehler werden alle Korrelationen um den gleichen Faktor erhöht, an den grundsätzlichen Zusammenhängen ändert sich auch bei suboptimaler Spezifikation dieses Faktors nichts.

Ein weiteres methodisches Problem, das mit Metaanalysen praktisch immer einhergeht, ist die relativ geringe Samplegröße. Mit 143 Fällen bewegt sich die Arbeit am oberen Rand vergleichbarer Studien, dennoch sind Limitationen unübersehbar, insbesondere, wenn Interaktionseffekte oder Gruppierungsvariablen mit vielen Ausprägungen (Branche, Leistungsoperationalisierung disaggregiert etc.) betrachtet werden sollen. Im Fall der Interaktionseffekte werden mehrere Gruppierungsvariablen hintereinander angewandt, das Sample der Metaanalyse also in mehreren Schritten unterteilt. Die Subgruppen, die hieraus resultieren, lassen statistisch belastbare Aussagen kaum mehr zu. Auch Variablen mit vielen Ausprägungen führen zu teils sehr kleinen Subgruppen, insbesondere dann, wenn die Ausprägungen unterschiedlich stark besetzt sind. Dieses Problem ist für Metaanalysen so nicht lösbar und verhindert an einigen Stellen weitergehende Analysen, es bleibt nur, an den entsprechenden Stellen auf die Begrenztheit der Aussagekraft hinzuweisen, wie in Abschnitt 5.4 dieser Arbeit jeweils geschehen. Das Sample ist auf der anderen Seite groß genug, um die meisten Gruppierungsvariablen mit befriedigenden resultierenden Gruppengrößen anwenden zu können. Hierdurch werden auch Berechnungen möglich, die Metaanalysen mit kleineren Samples so nicht durchführen konnten, etwa die Berechnung von Effektstärken.

5.6.2 Einordnung in den wissenschaftlichen Diskurs

Da sich die hier präsentierte Metaanalyse in Sample und teilweise auch einzelnen Forschungsfragen relativ eng an die Studie von Crook et al. (2008) anlehnt, soll im Interesse guter wissenschaftlicher Praxis noch eine Abgrenzung der beiden Arbeiten vorgenommen und die Metaanalyse in den bestehenden wissenschaftlichen Diskurs eingeordnet werden. Beide Arbeiten unterscheiden sich zunächst wesentlich im Erkenntnisinteresse: Anstatt den Zusammenhang zwischen Performance und Ressourcen im Gedankengebäude des Resource-based view zu untersuchen und hierbei methodische Artefakte zu eliminieren, das Anliegen des Artikels von Crook und Kollegen, untersuche ich exakt diese methodischen Artefakte. Ihr Ausmaß soll Aufschluss über die Bedeutung der im bisherigen Gang der Arbeit diskutierten methodischen Schwierigkeiten in der Strategischen Managementforschung geben.

Crook et al. (2008) betrachten ebenfalls die Auswirkungen appropriationsrobuster Kennzahlen, nutzen hierzu jedoch lediglich eine bivariate Variable (robust – nicht robust). Dies scheint aufgrund der im Rahmen der Literaturreview in Abschnitt 4.2.3 ermittelten Heterogenität in den Leistungsoperationalisierungen nicht befriedigend. Die vorliegende Untersuchung geht entsprechend einen Schritt weiter, um ein differenzierteres Bild des Einflusses des Appropriationsprozesses auf die Ergebnisse von Studien im Fachgebiet zu erlangen und auch die Stärke des Beeinflussungseffektes zu messen. Auch fehlt der Arbeit von Crook und Kollegen in weiten Teilen ein theoriegestütztes Erklärungsmodell für Appropriationsprobleme, wie es in der hier vorliegenden Arbeit in den Abschnitten 3.3 und 5.5.2 entwickelt wurde.

Die Untersuchung von Crook und Kollegen leidet überdies an in ihrem Versuch, Ressourcen homogen zu beschreiben, unter einem erheblichen „apples and oranges“-Problem (Drinkmann, 1990; vergleiche auch Abschnitt 5.1.2). Unter dem Begriff Ressourcen werden hier hochgradig unterschiedliche Konstrukte gefasst, deren Kausalität zur Performance in ganz anderen theoretischen und empirischen Diskursräumen als dem des Resource-based view diskutiert werden muss. So untersuchen etwa Perry-Smith & Blum (2000) den Zusammenhang zwischen Praktiken des Personalmanagements, die auf eine nachhaltige Work-Life-Balance der Mitarbeiter aufbauen vor dem Hintergrund des Resource-based view, wohingegen Li, Lam & Qian (2001) die Erfahrung des Managements internationaler Joint-Ventures in einem bestimmten Kulturkreis mit der Performance dieser Joint Ventures in Verbindung bringen und Richard (2000) die kulturelle Diversität der Belegschaft als Ressource im Sinne des Resource-based view interpretiert. Es erfordert ein erhebliches Abstraktionsvermögen, internationale Managementenerfahrung, Work-Life-Balance als Personalmanagementmaßnahme und Diversity als Ausprägung eines gemeinsamen Konstrukts „Ressource“ zu interpretieren. Eher entstammen diese Beziehungen und Forschungsfragen im Kern anderen Diskursräumen: Der Kultur-, HRM- und Diversity-Forschung. Methodisch ist dies insbesondere bedenklich, da sich manche Ressourcen (etwa Innovationsfähigkeit) kausal deutlich näher an der Performance bewegen als andere (etwa bestimmte HRM-Praktiken). Damit ergeben sich zwangsläufig höhere (beziehungsweise niedrigere) Korrelationskoeffizienten. Dies schränkt als maßgebliche unerklärte Variable die Aussagekraft der gewonnenen Erkenntnisse ein und stellt überdies eine weitere Moderatorvariable dar, die die Heterogenität der Forschungsergebnisse erklären kann. Diesen grundsätzlichen Vorbehalt gegen die Arbeit von Crook und Kollegen reduziere ich in meinem Ansatz durch einen deutlich konservativeren Aufbau des Samples, der durch eine wesentlich rigorosere Anwendung der Definition von Ressourcen im Sinne des Resource-based view (vergleiche Barney, 1991) zur Abgrenzung der in die Metaanalyse ein-

zubeziehenden Studien erreicht wird. Dies erreiche ich durch die Eliminierung von Primärstudien mit zweifelhaftem Ressourcenbezug in der erklärenden Variablen sowie solchen, die sich sehr nahe an der Tautologie bewegen. Insbesondere das Kriterium der Nichtimitierbarkeit der Ressource wird meiner Einschätzung nach relativ häufig verletzt. Die Eliminierung entsprechend fragwürdiger Arbeiten aus dem Sample ist eine Möglichkeit um eine Veredelung des Datensatzes und damit eine Verbesserung der Ergebnisse zu erreichen.

Schlussendlich wende ich tiefergehende statistische Methoden an, die einen weiteren Erkenntnisfortschritt ermöglichen. Die Arbeit von Crook und Kollegen basiert lediglich auf der Abschätzung der Unterschiedlichkeit von Varianzen und Mittelwerten, während im Fall der hier vorgelegten Arbeit durch die Anwendung von gesättigten ANOVA-Modellen weitergehende Analysen, etwa über Effektstärken und insbesondere Interaktionsbeziehungen zwischen erklärenden Variablen möglich werden. In Summe glaube ich die empirischen Untersuchungen zu systematischen Methodenproblemen in der Strategischen Managementforschung im Vergleich zur Arbeit von Crook und Kollegen deutlich weiterentwickelt zu haben.

Zusammenfassend lässt sich also feststellen, dass sich die hier vorgelegte Untersuchung in inhaltlicher (Chakravarthy, 1986; March & Sutton, 1997; Short, Ketchen & Palmer, 2002; Combs, Crook & Shook, 2005; Ketchen, Boyd & Bergh, 2008; Crook et al., 2008 und andere) und methodischer Sicht (Cooper & Hedges, 1994; Orlitzky, Schmidt & Rynes, 2003; Crook & Combs, 2007; Chang et al., 2009 und andere) in einen bestehenden Diskursraum einordnet, diesen aber an verschiedenen Stellen entscheidend weiterentwickelt. Hieraus bezieht sie ihre eigenständige wissenschaftliche Rechtfertigung.

6. Zusammenführung und Diskussion

6.1 Zusammenfassung der Kernergebnisse

6.1.1 Erkenntnisinteresse, Forschungsbedarf und Gang der Arbeit

Die Strategische Managementforschung wird sehr häufig für ihre Unfähigkeit kritisiert, konsistente Ergebnisse zu generieren. Zu sehr vielen Forschungsfragen liegt ein Strauß an Arbeiten vor, die, obgleich sie das gleiche Thema mit ähnlicher Methodik bearbeiten, doch zu vollständig gegensätzlichen Ergebnissen gelangen. Gründe hierfür zu diskutieren, diese deduktiv zu durchdringen und schließlich empirisch auf Relevanz zu überprüfen ist der Forschungsgegenstand der vorliegenden Arbeit.

Ziel war also, die Heterogenität von Ergebnissen in der Strategischen Managementforschung zu erklären. Es galt der Vermutung nachzugehen, dass die zentrale abhängige Variable, die Leistung einer Organisation, hierbei eine entscheidende Rolle spielen könnte. Ausgangspunkt dieser Vermutung war die Beobachtung, dass eine Reihe von Studien vorliegen, die mehrere Leistungsindikatoren als abhängige Variablen mit den selben unabhängigen Variablen in einen Zusammenhang bringen; der Zusammenhang stellt sich je nach verwendetem Leistungsindikator häufig ganz anders dar. Zugleich äußern eine Reihe Autoren Unbehagen ob des Wissensstandes um diese zentrale abhängige Variable im Fachgebiet (Chakravarthy, 1986; Lebas, 1995; Zimmerman, 2001; Orlitzky, Schmidt & Rynes, 2003; Powell, 2003b; Boyd et al., 2005; Combs et al., 2005; Richard et al., 2009). Neben der Operationalisierung der Leistung kommt auch der Abgrenzung der Teilpopulation, aus der die Stichprobe einer Studie gezogen wird, sowie dem unklaren zeitlichen Zusammenhang zwischen Ursache und Wirkung eine Schlüsselrolle bei der Analyse der Heterogenität der Studienergebnisse in der Strategischen Managementforschung zu.

Empirisch galt es, den state of the art in der Strategischen Managementforschung hinsichtlich der Operationalisierung von Leistung und des Umgangs mit den beiden anderen vermuteten Einflussfaktoren auf die Heterogenität von Studienergebnissen, die Populationsselektion und das Zeitversatzproblem, zu beschreiben, den Einfluss, den die drei genannten Methodenprobleme auf die Ergebnisse in der Strategischen Managementforschung ausüben, zu quantifizieren und so die deduktiv gewonnenen Annahmen in der Empirie zu überprüfen. Schließlich waren Empfehlungen für die wissenschaftliche Praxis abzuleiten.

6.1.2 Epistemologische Betrachtungen, der Leistungsbegriff und das Fachgebiet der Strategischen Managementforschung

Am Beginn der Betrachtungen stand eine Analyse der Art der Wissensgewinnung, vor deren Hintergrund die aufgeworfenen Fragen verstanden werden können. Verständnisleitend für diese Arbeit ist das kritisch-rationalistische Verständnis der Wissenschaften, wie es insbesondere von Gottfried Wilhelm Leibniz, Immanuel Kant und Karl Popper entwickelt wurde. Ausgehend von der Annahme der Existenz einer äußeren Realität findet Wissensfortschritt durch die Prüfung von Hypothesen statt, die umso gehalt- und damit für den Erkenntnisgewinn wertvoller sind, je mehr Fälle sie ausschließen. Diese Hypothesen wiederum werden aus Theorien gewonnen, die sich aus Sätzen zusammensetzen, die auf in sich schlüssigen Axiomensystemen basieren. Durch die stetige Weiter- oder Neuentwicklung von falsifizierten

Theorien findet Wissensfortschritt statt. Im Rahmen sozialwissenschaftlicher Arbeiten ist jedoch insbesondere die Zeit- und Kontextgebundenheit wissenschaftlicher Ergebnisse als prohibitiv für Verfahren anzusehen, die der „reinen Lehre“ Poppers, dem Verwerfen aller einmal falsifizierter Theorien, folgen wollen. Nötig und in der Praxis handlungsleitend ist die ständige Weiterentwicklung von Theorien, solange diese gute Erklärungsbeiträge für bestimmte Kontexte liefern und nicht durch bessere Theorien abgelöst wurden, auch wenn sie in bestimmten Zusammenhängen falsifiziert werden. Insbesondere vor dem Hintergrund der immer wieder beobachteten Heterogenität in den Ergebnissen von Studien zu gleichen Themen, dem Ausgangspunkt der Betrachtungen dieser Arbeit, ist dieses Vorgehen notwendig, da regelmäßig nicht entschieden werden kann, ob eine beobachtete Falsifizierung einer Hypothese tatsächlich auf den falsch vorhergesagten kausalen Zusammenhang zwischen erklärender und erklärter Variablen oder auf unbeobachtete Einflussfaktoren zurückzuführen ist. Die Entstehung von Wissen wird als Prozess der Deduktion, der Ableitung spezifischer Aussagen aus Allgemeinen, verstanden. Aus allgemeinen Theorien werden spezielle Hypothesen abgeleitet, die in bestimmten Kontexten und Erklärungszusammenhängen geprüft werden. Ergebnisse dieses Prüfprozesses fließen dann wieder in den Diskurs ein. Diese Logik des Vorgehens hat sich in den Wirtschaftswissenschaften bewährt und bis heute weitgehend durchgesetzt.

Hypothesentests basieren auf der Prüfung kausaler Aussagen, das heißt Aussagen darüber, ob und wie Umstand A zu Zustand B führt. Im Rahmen des in dieser Arbeit beleuchteten Themenfeldes der Strategischen Managementforschung ist der interessierende Zustand B sehr häufig die Leistung einer Organisation. Das Konstrukt der Leistung ist jedoch äußerst vielschichtig. Im Kontext der Strategischen Managementforschung findet, zumeist ohne dass dies expliziert würde, ein ergebnisorientierter Leistungsbegriff Verwendung. Das heißt Leistung ist nicht die Handlung des Leistens, sondern das Ergebnis einer Handlung, beziehungsweise das Ergebnis einer Serie von Handlungen. Für eine Gesamtorganisation ist die Leistung, so die Arbeitsdefinition der hier vorgelegten Studie, die Summe der Ergebnisse der Tätigkeiten ihrer Mitglieder. Woran dieses Ergebnis gemessen wird ist jedoch in den seltensten Fällen objektiv evident sondern immer Gegenstand subjektiver Entscheidungsprozesse des Messenden und überdies von den Umständen (insbesondere Zugang zu und Verfügbarkeit von Daten) bestimmt. Leistung ist zudem immer relativ, gewinnt ihren Sinn also erst durch den Vergleich mit einer anderen Größe, sei es einer Sollvorgabe, dem Vergleich mit anderen Organisationen oder einer Längsschnittbetrachtung, also dem Vergleich mit in der Vergangenheit liegenden Messpunkten. Um sinnvoll im Sinne der Strategischen Managementforschung verwendbar zu sein, muss der Indikator, der das Konstrukt Leistung beschreibt, außerdem quantifizierbar sein.

Die Perspektive eines quantitativ arbeitenden Wissenschaftlers, die dieser Arbeit zu Grunde liegt, unterscheidet sich dabei deutlich von der Perspektive auf die Leistung, die aus Sicht der Unternehmenssteuerung, dem Fachbereich des Business Performance Measurement, eingenommen wird. In letzterem Fall liegen in der Regel tiefgreifende Informationen aus der interessierenden Organisation vor, die Untersuchungsobjekte sind vorher bekannt, es sind relativ wenige, Messinstrumente können für den Einzelfall individuell ausgewählt oder entwickelt werden. Der quantitativ arbeitende Wissenschaftler dagegen ist an Zusammenhängen interessiert, die nicht nur für Einzelfälle gelten, er arbeitet daher in der Regel mit einer relativ großen Zahl an Organisationen. Diese Organisationen müssen allesamt mit demselben In-

strumentarium zur Messung der Leistung bewertet werden, um Vergleiche zu ermöglichen. Zumeist ist auch zu Beginn der Untersuchung nicht klar, welche Unternehmen beleuchtet werden sollen, im Gegenteil kann eine Zufallsauswahl oder gelungen gewählte Quota-Auswahl dieser Unternehmen aus einer potentiell sehr umfangreichen Grundgesamtheit als Qualitätsindikator verstanden werden. In aller Regel verfügt der Wissenschaftler darüber hinaus als den interessierenden Organisationen nicht Angehöriger über einen limitierten Zugang zu Informationen. Dies schränkt die Möglichkeiten im Vergleich zum Business Performance Measurement weiter ein. Allgemein gilt so, dass Indikatoren, die in der Strategischen Managementforschung Verwendung finden können, auch im Business Performance Measurement brauchbar sind, der Umkehrschluss gilt jedoch nicht. Komplexere Leistungsbeurteilungsinstrumente wie etwa Balanced Scorecards lassen sich im Kontext der Strategischen Managementforschung aus den oben genannten Gründen nicht anwenden. Die beiden Fachgebiete sind damit in Bezug auf die Operationalisierung von Leistung getrennt zu betrachten.

Schließlich wurde das hier interessierende Fachgebiet der Strategischen Managementforschung beleuchtet. Das zentrale Erkenntnisinteresse dieses Fachdiskurses liegt in der Erklärung mittel- und langfristiger Unterschiede in der Leistung von Organisationen und der Analyse des Verhaltens von Unternehmen auf einer strategischen Ebene. Die Strategische Managementforschung stellt damit ein Querschnittsfach dar, Überschneidungen zur Organisationssoziologie und -psychologie, zur Mikroökonomie oder zu den Politikwissenschaften sind umfangreich gegeben. Methodisch abzugrenzen ist der Ansatz der Strategischen Managementforschung von der Organisationskybernetik. Letztere versucht, Komplexität in den Zusammenhängen möglichst vollständig zu erfassen, während die Strategische Managementforschung reduktionistisch vorgeht, also versucht, komplexitätserhöhende Faktoren so weit wie möglich auszublenden und damit *ceteris paribus*-Situationen zu schaffen. Beide Ansätze scheinen inkommensurabel, die derzeitige empirische Forschung dominiert eindeutig die komplexitätsreduzierende Variante, die auch für die vorliegende Arbeit verständnisleitend ist.

6.1.3 Konzeptionelle Ansätze zur Erklärung der Ergebnisheterogenität in der Strategischen Managementforschung

Im Anschluss an diese grundlegenden Betrachtungen wurden drei mögliche Ursachen für die beobachtete Heterogenität in den Ergebnissen von Arbeiten der Strategischen Managementforschung konzeptionell analysiert: das Problem der Verallgemeinerbarkeit von Forschungsergebnissen über unterschiedliche Unternehmenspopulationen, das Problem der Wertschöpfung und -verteilung zwischen Stakeholdergruppen in Organisationen sowie das Problem der unklaren zeitlichen Zusammenhänge zwischen Ursache und Wirkung.¹³⁵

Das Problem der Populationsselektion tritt auf, da die Grundgesamtheit der Strategischen Managementforschung, die Gesamtheit aller Unternehmen zu jedem beliebigen Zeitpunkt, unüberschaubar groß ist. Sie verschließt sich so einer Vollerhebung, auch eine Stichprobenziehung mit Quota-Verfahren ist aufgrund der Heterogenität dieser Grundgesamtheit kaum möglich. Entsprechend ist es notwendig, die Grundgesamtheit auf eine handhabbare Teilpo-

¹³⁵ Aus Gründen der Fokussierung und Argumentationsklarheit konzentriert sich der Hauptteil der Arbeit auf diese drei zentralen Methodenprobleme, eine Reihe weiterer in der Literatur diskutierter Problemfelder wird in Anhang A.3.2 thematisiert und eingeordnet.

pulation einzugrenzen, etwa indem lediglich Unternehmen einer bestimmten Branche in einer einzelnen Volkswirtschaft zu einem definierten Zeitpunkt betrachtet werden. Mit der Auswahl dieser Teilpopulation, die sehr häufig nicht aus der Logik der zu bearbeitenden Forschungsfrage sondern der des Forschungsprozesses gewählt wird, geht ein Verlust an Verallgemeinerbarkeit einher. Die betrachtete Teilpopulation ist nicht mehr notwendigerweise repräsentativ für die Grundgesamtheit, für die eine Aussage intendiert und unter Umständen auch suggeriert wird. Das Erkenntnisinteresse von Arbeiten der Strategischen Managementforschung ist in der Regel ja die Aufdeckung generalistisch gültiger Zusammenhänge. Die mit der Teilpopulation einhergehende Begrenzung der Aussagekraft der Forschungsergebnisse wird in der Forschungspraxis häufig nicht weiter diskutiert, spätestens bei der Zitation der Ergebnisse in anderen Studien wird Forschungsergebnissen eine Aussagekraft unterstellt, die aufgrund der begrenzten Größe der untersuchten Teilpopulation so nicht gegeben ist. Es kommt zu einer Entkontextualisierung von Forschungsergebnissen. Folgerichtig ergeben sich, wenn unterschiedliche Studien gleiche Zusammenhänge in unterschiedlichen Teilpopulationen betrachten, heterogene Ergebnisse. Am häufigsten wird die Grundgesamtheit aller Unternehmen nach Land, Branche, Notierung und Größe selektiert. Aus anderen Forschungen liegen Erkenntnisse vor, die die Relevanz dieser Selektionsfilter nahelegen, die also annehmen lassen, dass sich kausale Zusammenhänge in unterschiedlichen Branchen, für Unternehmen unterschiedlicher Größe und so weiter anders darstellen könnten. Wäre dem so, wäre die Branche, Größenklasse und so weiter, aus der Unternehmen für eine bestimmte Studie gezogen werden, ein möglicher moderierender Einfluss auf Forschungsergebnisse in der Strategischen Managementforschung.

Ein vielversprechender Ansatzpunkt zum Verständnis des Einflusses der Leistungsoperationalisierung auf die Ergebnisse der Strategischen Managementforschung bietet die Appropriationstheorie. Diese beschreibt die Verteilung der Wertschöpfung in einer Organisation zwischen Stakeholdergruppen. Ausgehend von der Kernperformance, die die gesamten geschaffenen Werte umfasst, versuchen alle Stakeholdergruppen einer Organisation, von Zulieferern und Kunden über Mitarbeiter und Management bis hin zu den Eigentümern, einen möglichst großen Teil dieser Wertschöpfung zu vereinnahmen. Die Verteilung der geschaffenen Werte ist damit ein politischer Prozess, dessen Ausgang durch die relative Verhandlungsmacht der Stakeholdergruppen bestimmt wird. Je nachdem wo in diesem Prozess die Leistung gemessen wird sind mehr oder weniger unerklärte Faktoren in einem beobachteten Leistungsergebnis beinhaltet, nämlich die Verhandlungsmacht von keinen, wenigen oder vielen Stakeholdergruppen. Arbeiten Studien mit Leistungsoperationalisierungen, die nah am Konzept der Kernperformance messen, sollten sie, dieser Theorie zufolge, stärkere Zusammenhänge zwischen leistungsbeeinflussenden Faktoren und der Leistung der untersuchten Unternehmen finden. Demgemäß sollten Arbeiten, die die Leistung mittels markt- und accountingbasierte Kennzahlen erfassen, deutlich schwächere Beziehungen zwischen Leistung und erklärenden Variablen finden als solche, die die Unternehmensperformance prozessbasiert und damit kernwertschöpfungsnah messen.

Diese Fragestellung ist für die Strategische Managementforschung jedoch nur dann von Relevanz, wenn tatsächlich ein Auswahlproblem hinsichtlich der Operationalisierung von Leistung besteht, es also nicht eine logische oder dominante Leistungsoperationalisierung gibt, die als Maßstab fungiert. Folgt man der Agency-Theorie gibt es eine solche, nämlich die Appropriationsleistung der Prinzipale der Organisation, der Anteilseigner. Ich argumentierte dagegen,

dass die Agency-Theorie für die Fragestellungen der Strategischen Managementforschung nicht verständnisleitend sein kann und begründe diese Annahme durch eine Analyse des Organisationsverständnisses der Agency-Theorie, die die Organisation als Nexus expliziter Verträge versteht. Durch die Annahme, dass die Teilnehmer der Organisation als ihren Individualnutzen maximierende Agenten nur so viel in die Organisation einbringen wie sie zwingend, also durch einen expliziten Kontrakt vorgeschrieben, müssen, lässt sich die Leistung einer Organisation vollständig als die Summe der vertraglich festgeschriebenen Leistungen der Teilnehmer beschreiben. Leistungen müssen in diesem Organisationsverständnis zu Marktpreisen eingekauft werden, Heterogenität in der Performance von Organisationen, das zentrale Erkenntnisinteresse der Strategischen Managementforschung, ist mit diesem Modell nicht abbildbar. Notwendig ist also eine andere Basistheorie, die Organisationstheorie der expliziten und impliziten Verträge. Diese erlaubt die Abkehr von der Annahme rein eigennutzmaximierenden Verhaltens und damit die Abbildung von Performanceunterschieden zwischen Organisationen. Allerdings verliert in dieser Organisationstheorie die logische Zielfunktion ihre Fundierung, es gibt keine Prinzipale mehr, vielmehr ringen verschiedene Stakeholder um Einfluss. Die Definition einer Zielfunktion ist nun nicht mehr von außen, durch die Basistheorie, vorgegeben sondern vielmehr Aufgabe der jeweiligen Entscheider, im Zuge der Unternehmenssteuerung ist dies die Unternehmensführung, für Überlegungen zur wissenschaftlichen Analyse der Leistung von Organisationen sind es die durchführenden Wissenschaftler. Folgt man dieser Argumentation kann es durch den Wegfall dieser eindeutigen Zielfunktion keine aus theoretischen Überlegungen heraus überlegene Leistungsoperationalisierung geben. Vielmehr muss diese so gewählt werden, dass sie für die zu untersuchenden Forschungsfrage funktional und angemessen ist. Unter dieser Annahme ist die Appropriationstheorie relevant. Die Möglichkeit, Zusammenhänge zwischen erklärenden Variablen und der Performance abzubilden könnte, bei Betrachtung der selben erklärenden Variablen, wesentlich von der Nähe der abhängigen Variablen „Leistung“ zur Kernperformance abhängen. Dies könnte einen Teil der Heterogenität der Studienergebnisse in der Strategischen Managementforschung erklären.

Das dritte diskutierte Problem betrifft den unklaren zeitlichen Zusammenhang zwischen erklärenden und erklärten Variablen. Insbesondere in der Strategischen Managementforschung, die darauf angewiesen ist, sehr komplexe Zusammenhänge auf einfache und überprüfbare Kausalbeziehungen zu reduzieren, ist das Wissen zentral, wann eine bestimmte Aktion zu einer zugehörigen Reaktion führt. Im Extremfall kann es zum Problem der reversed causalities kommen, wenn tatsächlich die abhängige Variable eine entsprechende Veränderung der unabhängigen bewirkt anstatt umgekehrt. Grundsätzlich ist davon auszugehen, dass, wenn die Gefahr besteht die abhängige Variable zum falschen Zeitpunkt zu messen, Zusammenhänge tendenziell verschleiert werden. Es stehen verschiedene Instrumentarien zur Verfügung, um den Zeitversatz zwischen Ursache und Wirkung zu modellieren. So erlauben etwa Paneldatenuntersuchungen die Einbeziehung verschiedener Messzeitpunkte für die abhängige Variable, mit Dummy-Variablen kann für den Einfluss eines bestimmten Zeitpunktes kontrolliert werden, Experten wird die Fähigkeit zugeschrieben, Entwicklungen der näheren Zukunft zu prognostizieren, auch die Verwendung von über längere Perioden gemittelten Leistungsindikatoren stellt einen Weg zur Eingrenzung des Problems dar. Die aus den konzeptionellen Überlegungen heraus empfehlenswerteste Vorgehensweise ist die deduktive Ableitung von Hypothesen über den erwarteten zeitlichen Zusammenhang zwischen abhängigen und

unabhängigen Variablen sowie dessen unmittelbare Überprüfung im empirischen Modell. Die Berücksichtigung oder Nichtberücksichtigung des Zeitversatzproblems könnte in jedem Fall eine unbeobachtete Variable darstellen, die die Heterogenität von Forschungsergebnissen im Fachgebiet erklären könnte.

6.1.4 Empirische Analyse

Am Beginn des folgenden empirischen Teils der Arbeit stand zunächst eine qualitative Literaturreview. Ziel war die Ermittlung eines Status Quo der Strategischen Managementforschung. Hierzu wurden drei Jahrgänge der beiden führenden Journals in der Strategischen Managementforschung, des Strategic Management Journal und des Academy of Management Journal, vollständig erhoben. Die zu überprüfenden Forschungsfragen sind der nachfolgenden Tabelle zu entnehmen.

Tabelle 40: Forschungsfragen der qualitativen Literaturreview

Nr.	Forschungsfrage
F1	<i>Wie relevant ist die Operationalisierung von Leistung für die Strategische Managementforschung?</i>
F2	<i>Findet sich ein dominantes Paradigma der Leistungsmessung in der Strategischen Managementforschung?</i>
F3	<i>Nach welchen Kriterien werden Teilpopulationen abgegrenzt? Wird das Problem der begrenzten Aussagekraft, die durch die Verwendung begrenzter Teilpopulationen zu Stande kommt, in Arbeiten der Strategischen Managementforschung erkannt und benannt?</i>
F4	<i>Wird das Problem des Zeitversatzes erkannt? Welche Verfahren werden angewandt, um es einzugrenzen?</i>

Mehr als die Hälfte aller Arbeiten, die in den beiden Journals erschienen sind, legen Untersuchungsdesigns zu Grunde, die eine Operationalisierung von Leistung voraussetzen. Die Operationalisierung von Leistung ist im Fachgebiet also in der Tat zentral. Jeweils etwa ein Viertel dieser Studien arbeitet mit accounting-, markt- und prozessbasierten Kennzahlen, die übrigen 25 Prozent verteilen sich auf expertenbasierte und feldspezifische Operationalisierungen, ein dominantes Paradigma der Leistungsoperationalisierung ist also nicht zu beobachten. Dies bestätigt die Überlegungen zur Agency-Theorie, diese scheint in der Forschungspraxis nicht verständnisleitend zu sein. Insgesamt finden 27 verschiedene Indikatoren zur Leistungsmessung Verwendung, am häufigsten werden Kapitalrentabilitäten, Einschätzungen interner Experten, Market-to-Book Ratios, Umsatz und Überleben genutzt. Im Schnitt arbeiten die Studien mit 1,3 Performanceoperationalisierungen, es kommt also relativ häufig vor, dass die Leistung in einer Untersuchung mit zwei oder mehr unterschiedlichen Indikatoren erfasst wird. Angesichts der Möglichkeit, dass unterschiedliche Leistungsoperationalisierungen zu unterschiedlichen Studienergebnissen führen können, ist dies zu begrüßen. Teilpopulationen werden insbesondere über Branchen, Volkswirtschaften sowie Größe und Notierungsstatus der einbezogenen Unternehmen abgegrenzt, trivialerweise notwendig ist überdies die Beschränkung der Arbeit auf einen bestimmten Zeitraum. Das Problem des Zeitversatzes wird, anders

als von anderen Autoren noch befürchtet (Bergh, Hanke, Balkundi, Brown & Chen, 2004, Greve & Goldeng, 2004), in der aktuellen Forschung erkannt und breit rezipiert. Nur eine Minderheit der in der qualitativen Literaturreview untersuchten Studien arbeitet mit Verfahren, die für Zeitversatzprobleme anfällig sind.

Die drei deduktiv ermittelten Ansätze zur Erklärung der Heterogenität von Ergebnissen in der Strategischen Managementforschung (Populationsselektion, Appropriation von Wertschöpfung, Time Lag) galt es im Anschluss auf statistische Relevanz zu überprüfen. Gewählt wurde hierzu der Ansatz einer Metaanalyse, die die Varianz der Ergebnisse von Primärstudien in einem häufig beforschten Teilgebiet der Strategischen Managementforschung, dem Zusammenhang zwischen Performance und Ressourcen im Sinne des Resource-based view, analysieren und mit den drei diskutierten methodischen Problemen in Verbindung bringen sollte. Hierzu wurde ein Datensatz mit publizierten Arbeiten zum Zusammenhang zwischen Ressourcen und Performance erstellt. Die in den jeweiligen Arbeiten gefundenen Stärken dieses Zusammenhangs wurden in Form der jeweils ermittelten Korrelationskoeffizienten als abhängige Variable der Metaanalyse definiert. Aus den deduktiven Überlegungen heraus wurden Gruppierungsvariablen erstellt, die den Datensatz der Metaanalyse anhand interessierender Kriterien der jeweiligen Primärstudien aufteilen.

Zentrales Problem metaanalytischer Untersuchungen ist die in der Regel recht begrenzte Größe des erstellbaren Datensatzes. Da zu vielen Zusammenhängen nur relativ wenige Primärstudien vorliegen, ist der mögliche Erkenntnisgewinn häufig gering. Notwendig zur Beantwortung der Forschungsfragen der hier vorgelegten Arbeit war daher einen Zusammenhang zu analysieren, zu dem eine ausreichend große Zahl an Primärstudien vorliegt. Hier bot sich die Verwendung des Datensatzes zum Resource-based view an, da zum Zusammenhang zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen eine Anzahl an Studien vorliegt, die im Rahmen der Metaanalyse auch aufwändigere statistische Verfahren anwendbar macht. Überdies ist nach dem bisherigen Kenntnisstand begründet davon auszugehen, dass tatsächlich eine kausale Verknüpfung zwischen abhängiger und erklärender Variablen besteht, die durch unbeobachtete beeinflussende Variablen, etwa die Art der Leistungsoperationalisierung, mehr oder weniger stark verschleiert wird. Auch dies ist eine Voraussetzung der durchgeführten Metaanalyse.

Folgende Tabelle fasst die deduktiv gewonnenen und in der Metaanalyse getesteten Hypothesen sowie die erzielten Ergebnisse der Hypothesentests zusammen:

Tabelle 41: Zusammenfassung Hypothesen Metaanalyse

Nr.	Hypothese	Ergebnis
H1.1	<i>Korrelationen zwischen Leistung und Ressourcen unterscheiden sich signifikant, je nachdem, welche Branchen betrachtet werden.</i>	nicht bestätigt
H1.2	<i>Studien, die mehr als eine Branche umfassen, berichten geringere Korrelationen und weisen höhere Varianzen auf als Studien, die sich auf eine oder wenige Branchen konzentrieren.</i>	nicht bestätigt
H1.3	<i>Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, welche Zeitabschnitte in den zu Grunde liegenden Primärstudien betrachtet werden.</i>	Keine Aussage möglich, von H2 überlagert

H1.4	<i>Korrelationen unterscheiden sich in statistisch signifikanter Weise, je nachdem, ob gelistete oder nicht gelistete Unternehmen untersucht werden.</i>	Keine Aussage möglich, von H2 überlagert
H2	<i>Die Korrelation zwischen Performance und Ressource fällt schwächer aus, je weiter sich die zu Grunde gelegte Performanceoperationalisierung vom Konzept der Kernperformance entfernt, je mehr sie also bereits Objekt des Verteilungskampfes zwischen Stakeholdergruppen war.</i>	bestätigt
H3.1	<i>Studien, die mit Time-Lag-robusten Leistungsoperationalisierungen arbeiten, finden einen stärkeren Zusammenhang zwischen Ressourcen und Performance als Studien, deren Performanceoperationalisierungen nicht robust gegenüber Zeitversatzphänomenen sind.</i>	Teilweise bestätigt
H3.2	<i>Der Teil der Studien, der nicht mit für Zeitversatz robusten Verfahren arbeitet, weist eine höhere Varianz in der abhängigen Variablen auf als der Teil, der robuste Verfahren einsetzt.</i>	nicht bestätigt

Im Rahmen der Analyse des Einflusses unterschiedlicher Teilpopulationen (Hypothese H1.1 bis H1.4) waren keine eindeutigen Ergebnisse zu erzielen. Während für unterschiedliche Branchen kein signifikanter Effekt auf die Stärke der Korrelation zwischen Performance und Ressourcen festgestellt werden konnte, lag für die anderen beiden Selektionsfilter, Zeitraum und Börsennotierung, ein statistisch signifikanter Einfluss vor. Dieser war jedoch zu großen Teilen auf die Beziehung der jeweiligen Populationsselektionsfilter zu bestimmten Leistungsoperationalisierungen zurückzuführen. Die Leistungsoperationalisierung überstrahlte in der durchgeführten Metaanalyse etwaige Einflüsse der Populationsselektion. Dies bedeutet im Umkehrschluss jedoch nicht, dass ein Einfluss der Populationsselektion auf Studienergebnisse nicht vorhanden wäre, sehr wohl jedoch, dass dieser Einfluss wesentlich geringer ausfällt als derjenige der Leistungsoperationalisierung. Auch die Relevanz des Time-Lag Problems kann nur teilweise bestätigt werden. Ein Effekt ist beobachtbar, er fällt aber sehr schwach aus. Studien, die mit Verfahren arbeiten, die für Zeitversatzprobleme robust sind, berichten überzufällig stärkere Korrelationen zwischen Ressource und Performance als Studien, die Ursache und Wirkung in der gleichen Periode messen, allerdings nur, wenn accounting- und markt-basierte Leistungsoperationalisierungen Verwendung finden. Hypothese H3.2 bestätigte sich nicht, die Berücksichtigung von Zeitversatzproblemen hat keinen überzufälligen Einfluss auf die Varianz der Ergebnisse in den durch die Gruppierungsvariable gebildeten Teilgruppen.

Den bei weitem größten Beitrag zur Aufklärung der Varianz der Studienergebnisse leistete die Appropriationstheorie und die aus ihr abgeleitete Hypothese H2. Es konnte klar gezeigt werden, dass, wie aus der theoretischen Modellierung heraus zu erwarten, markt-basierte Operationalisierungen die geringsten Korrelationen zwischen Ressource und Leistung anzeigen, gefolgt von accountingbasierten Indikatoren. Prozessbasierten Leistungsoperationalisierungen bilden klar stärkere Zusammenhänge ab. Die Vorhersagen der Appropriationstheorie werden also voll bestätigt, überdies erweist sich dieser Effekt als deutlich stärker als die Auswirkungen der beiden anderen Methodenprobleme. Auffällig ist, dass die Verwendung von

marktbasierten Leistungsoperationalisierungen zu noch schwächeren Korrelationen zwischen Leistung und Ressource führen als dies bei accountingbasierten Leistungsindikatoren der Fall ist. Dies deutet darauf hin, dass neben der in der Arbeit diskutierten Appropriationsproblematik hier weitere unbeobachtete Einflüsse vorliegen, etwa generelle Markt- und Börsentrends, die, unabhängig von der tatsächlichen wirtschaftlichen Situation eines Unternehmens, den Marktwert positiv oder negativ beeinflussen können.

Überraschend war der Einfluss der expertenbasierten Leistungsoperationalisierungen. Konzeptionell war vermutet worden, dass Experten, die in der aktuellen Forschungspraxis in der Regel Teil des Top-Managements der interessierenden Unternehmen sind, eine geistige Nähe zu accountingbasierten Kennzahlen aufweisen könnten. Entsprechend wären bei expertenbasierten Leistungsindikatoren Korrelationskoeffizienten der Primärstudien in etwa auf dem Niveau der accountingbasierten Indikatoren zu vermuten gewesen. Dies ist keineswegs der Fall, im Gegenteil berichten Arbeiten, die die Leistung von Unternehmen über die Einschätzung von Experten operationalisieren, die klar stärksten Zusammenhänge zwischen Ressource und Unternehmensleistung. Verschiedene Erklärungen sind hierfür denkbar, die wahrscheinlichste hierbei ist, dass Experten durch ihre informational privilegierte Stellung in der Lage sein könnten, von unbeobachteten Moderatorvariablen wie insbesondere dem Prozess der Verteilung der geschaffenen Werte in der Organisation zu abstrahieren. Diese Vermutung kann sich konzeptionell sowohl auf die Expertenforschung, als auch auf die Agency- und Appropriationstheorie stützen, Experten wird hier jeweils ein besonders guter Zugang zu Informationen unterstellt, forschungspraktisch ist dies ja auch ein wesentlicher Grund, um die Leistung von Experten einschätzen zu lassen anstatt auf zumeist leichter zugängliche accounting- oder marktbasierter Daten zuzugreifen.

Damit sind die Kernthesen der Arbeit bestätigt. Die Heterogenität der Ergebnisse in der Strategischen Managementforschung lässt sich tatsächlich mit unbeobachteten Variablen erklären. Zentral ist hierbei die Operationalisierung der Leistung eines Unternehmens. Die Implikationen der in der vorgelegten Arbeit gewonnenen Erkenntnisse sind im Folgenden zu diskutieren.

6.2 Diskussion und Ausblick

6.2.1 Diskussion und Implikationen für die wissenschaftliche Praxis

6.2.1.1 Heterogenität der Ergebnisse in der Strategischen Managementforschung

Im Rahmen dieser Arbeit wurde aufgezeigt, dass sich in der Tat ein Teil der Heterogenität der Ergebnisse in der Strategischen Managementforschung systematisch durch die angewandten Forschungsmethoden erklären lässt. Dieser Teil ist nicht gering, aber auch nicht so groß, als dass die Annahme von Hunter & Schmidt (2007) bestätigt werden könnte, die argumentiert hatten, dass Methodenprobleme nahezu die gesamte Varianz der Ergebnisse in einer Vielzahl von Forschungsfeldern erklären könnten und daher in einer sauberen Anwendung des Instrumentariums der Metaanalyse der Schlüssel zur Etablierung belastbarer Kausalitäten liegen könnte. Nichtsdestotrotz konnte gezeigt werden, dass unklare zeitliche Zusammenhänge zwischen erklärenden und erklärten Variablen, die Selektion von Teilpopulationen und insbesondere und vor allem die Operationalisierung von Leistung in der Tat die Ergebnisse von

Primärstudien in der Strategischen Managementforschung substantiell beeinflussen. Die Appropriationstheorie bietet eine schlüssige Erklärung für den beobachteten Einfluss der Leistungsoperationalisierung. Die Implikationen für die normative Aufgabe des Fachgebiets (vergleiche Abschnitt 2.3) sind erheblich: Je nach angewandter Methodik ändern sich die aus der Wissenschaft für die Praxis abgeleiteten Empfehlungen teilweise vollständig. Beim Rückgriff auf existentes Wissen in Form der Zitation vorliegender Studien oder der Ableitung von Empfehlungen für die Praxis ist also Vorsicht geboten. Einzelstudien sind mit relativ hoher Wahrscheinlichkeit für sich nur beschränkt aussagekräftig, metaanalytische Methoden bieten sich zur Zusammenführung von mit unterschiedlichen Methoden arbeitenden Studien zur Etablierung von Kausalaussagen an.

Wie in den Abschnitten 3.1 und 3.2 theoretisch und forschungspraktisch beleuchtet ist es kaum realistisch, die Gesamtheit aller Unternehmen als eine statistische Grundgesamtheit für Untersuchungen im Kontext der Strategischen Managementforschung zu verstehen, auf zu vielen Ebenen sind systematische Unterschiede zwischen Teilgruppen von Unternehmen zu erwarten. Grundsätzlich bieten sich zwei Möglichkeiten an, mit dieser Tatsache umzugehen. Zum einen lässt sich Bambergers Vorschlag aufgreifen (Bamberger, 2008), der eine Rekontextualisierung des Fachgebietes fordert. Dies würde bedeuten, dass Studien den Anspruch aufgeben, Zusammenhänge in der Grundgesamtheit aller Unternehmen erklären zu wollen. Zu Ende gedacht wäre es also beispielsweise nicht mehr möglich, eine Theorie über etwa Zusammenhänge zwischen HPWS und der Leistung von Organisationen zu entwickeln, vielmehr müsste eine solche Theorie spezifisch kontextgebunden erarbeitet werden, beziehungsweise umfangreiche Kontextdeterminanten einbeziehen. Die zweite und vielleicht befriedigendere Möglichkeit liegt in der Synthese von Forschungsergebnissen über Metaanalysen. So ist es möglich, ein Gesamtbild eines Zusammenhangs zu entwickeln und zumindest grob zu verstehen, ob ein Theoriegebäude in unterschiedlichen Kontexten tragfähig ist oder nicht. Auch könnten so Einflussfaktoren und Kontextdeterminanten gezielt erarbeitet werden. Durch die Einbeziehung von Moderatorvariablen lassen sich so auch die Stärken der Einflüsse bestimmter Kontextfaktoren ermitteln. Notwendig hierfür ist das Vorliegen einer befriedigend großen Zahl an Primärstudien zum selben Zusammenhang in unterschiedlichen Kontexten, wie auch von Singh et al. (2003) gefordert. Die Primärstudien müssen überdies auf einer sauberen theoretischen Modellierung zu prüfender Zusammenhänge, also einem deduktiven Forschungsansatz, basieren. Für eine Weiterentwicklung von Theorien ist demnach zu empfehlen, die Ergebnisse einzelner Studien rekontextualisiert zu betrachten, also im Verständnis des Samples, das der jeweiligen Studie zu Grunde liegt, und generelle Aussagen und Empfehlungen lediglich auf Basis einer Wissenssynthese in Form von Meta-Studien abzuleiten. Verfasser von Primärstudien sind entsprechend zu Bescheidenheit aufgerufen; wie sich im Rahmen der Literaturreview in Kapitel 4 gezeigt hat versäumen es jedoch selbst in qualitativ hochwertigen Journalen publizierende Autoren regelmäßig, die Begrenztheit der Aussagekraft der von ihnen erzielten Ergebnisse zu würdigen.

Fraglich ist, inwieweit sich die Aussagekraft von Studienergebnissen durch eine echte Zufallsauswahl der in Primärstudien einbezogenen Unternehmensstichprobe verbessern ließe. Eine solche findet, wie ebenfalls in Kapitel 4 gezeigt wurde, kaum jemals Verwendung. Grundsätzlich bietet eine ausreichend große Zufallsauswahl eine gute Möglichkeit, sich der Repräsentativität der gezogenen Stichprobe für eine größere Grundgesamtheit anzunähern (vergleiche zur Theorie der Stichprobenziehung beispielsweise Hartung et al., 2009). Im hier vorliegenden

Fachgebiet ist die Grundgesamtheit aller Unternehmen jedoch so groß, dass eine Zufallsauswahl, die ungebiaste Ergebnisse generieren möchte, mit nicht handhabbar großen Zufallsamples und zahlreichen verschiedenen Zugängen zu Daten umgehen müsste. Auf diese Weise dürfte das Problem der Populationsselektion daher kaum zu überwinden sein. Gleiches gilt für Quota-Verfahren, die die gesamte Grundgesamtheit repräsentieren sollen. An der Arbeit mit mehr oder weniger umfangreichen Teilpopulation wird also auch zukünftig kein Weg vorbei gehen.

Weitere Forschung ist zur Bewertung der Relevanz des Selektionsfilters Zeitraum nötig. In den Ergebnissen der vorgelegten Arbeiten deuten einige Indizien darauf hin, dass sich Kausalitäten über den Zeitverlauf tatsächlich ändern könnten. Zwar ist die empirische Evidenz aufgrund der relativ geringen Fallzahlen in den einzelnen Zeitfenstern nur begrenzt belastbar, doch zeigt sich durchaus, dass sich verändernde Umweltbedingungen auch die Erfolgsdeterminanten von Organisationen verändern können. Damit werden ähnliche Ergebnisse in anderen Studien (Miller & Shamsie, 1996; Kronborg & Thomsen, 2009) ansatzweise bestätigt. Zu klären bleibt, ob sich damit eine Art natürlicher Verfall der Erkenntnisse in der Strategischen Managementforschung ergibt oder ob bestimmte Grundmuster über die Zeit konstant bleiben, hier besteht Forschungsbedarf. Festzuhalten bleibt, dass der Einfluss, den die Populationsselektion auf Studienergebnisse ausübt, bei weitem geringer ist als der Einfluss der Operationalisierung der zentralen abhängigen Variablen „Leistung“. In Folgeforschungen wird es notwendig sein, die Leistungsmessung aus der Untersuchung zu eliminieren, um eine Überlagerung der Populationsselektionseffekte durch die Frage der Leistungsoperationalisierung zu verhindern und so den originär der Populationsselektion zuzuschreibenden Anteil der Heterogenität von Studienergebnissen bewerten zu können. Der vorliegende Datensatz war nicht umfangreich genug um dies leisten zu können.

In vollständiger Würdigung der Ergebnisse scheint die Befürchtung, dass die unklaren zeitlichen Zusammenhänge zwischen unabhängigen und abhängigen Variablen ein bedeutendes Problem in der Strategischen Managementforschung darstellen könnten, unberechtigt. Zum Einen ist dieses Problem, wie die Literaturreview zeigte, erkannt und findet Beachtung, zum Zweiten liegen verschiedene Möglichkeiten vor, mit dieser Schwierigkeit effektiv umzugehen. Insbesondere die Verwendung von Längsschnittdatensätzen bietet hier alle Optionen (vergleiche auch Greve & Goldeng, 2004; Certo & Semadeni, 2006). Schließlich ist der beobachtbare Einfluss auf Studienergebnisse sehr begrenzt.

Von den drei betrachteten Methodenproblemen übt das in der methodenkritischen Literatur zur Metaanalyse (vergleiche Abschnitt 2.3) noch am wenigsten systematisch beleuchtete, nämlich die Operationalisierung der abhängigen Variablen Leistung, den bei weitem größten Einfluss auf die Ergebnisse empirischer Arbeiten in der Strategischen Managementforschung aus. Die etwa von Combs et al. (2005) und Powell (2003b) erhobene Forderung nach einer verstärkten Beachtung dieser zentralen Operationalisierung in der Strategischen Managementforschung lässt sich aus den Erkenntnissen dieser Arbeit heraus daher nur noch einmal deutlich unterstreichen.

6.2.1.2 Die Operationalisierung von Leistung

Wie im Zuge der vorgelegten Arbeit klar wurde, kann es keine ideale Operationalisierung von Leistung geben. Unternehmen sind komplexe Organisationen mit heterogenen Zielfunktionen, belastbare theoretisch-konzeptionelle Fundierungen zur Untermauerung der konzeptionellen Dominanz einer bestimmten dieser Zielfunktionen und damit einer bestimmten Leistungsoperationalisierung liegen derzeit nicht vor. Dennoch kommt der Operationalisierung von Leistung eine enorme Bedeutung für den Aussagegehalt wissenschaftlicher Studien insbesondere im Bereich der Strategischen Managementforschung zu. Wie gezeigt werden konnte, unterscheiden sich Studienergebnisse grundlegend, abhängig insbesondere davon, wie die abhängige Variable Leistung gemessen wird. Der Einfluss der Leistungsoperationalisierung auf die Ergebnisse von Arbeiten in der Strategischen Managementforschung ist weit stärker als der bekannter weiterer Schwierigkeiten, wie der Abgrenzung der zu untersuchenden Grundgesamtheit und der unklaren zeitlichen Zusammenhänge zwischen abhängigen und erklärenden Variablen.

Aus Sicht der in dieser Arbeit diskutierten Appropriationstheorie ist Heterogenität in den Leistungsoperationalisierungen auch keineswegs ein Fehler sondern eine wesentliche methodische Stellschraube, um Zusammenhänge in komplexen Kontexten sinnvoll abbilden zu können. Zentral ist ausschließlich die Erkenntnis, dass es ein einheitliches Konstrukt „Leistung“ so schlicht nicht gibt. Vielmehr ist zwischen der Leistungserstellung, der Kernperformance in der Terminologie meiner Arbeit, und der Leistungsappropriation durch verschiedene Stakeholdergruppen zu unterscheiden.

Im Rahmen theoretischer Betrachtungen kommt auch Powell (2003a) zu dem Schluss, dass die Analyse eines so heterogenen Forschungsfeldes wie der Strategischen Managementforschung, das mit derart komplexen Entitäten wie großen Organisationen und ihren internen und externen Umwelten umgehen muss, sich notwendigerweise unterschiedliche Zugänge zur abhängigen Variablen offen halten muss. Powell argumentiert in einem anderen Beitrag (Powell, 2001) darüber hinaus auch gegen ein Suchen nach einer idealen Operationalisierung: „our task as scholars is not perfection or non-contradiction, but intellectual progress“ (S. 885). Ein solcher intellektueller Fortschritt allerdings ist auf der anderen Seite nur mit einem tiefen Verständnis dessen möglich, was als wesentliche erklärende Variable maßgeblich Design und Ergebnis von Studien bestimmt. Eine tiefgreifende Auseinandersetzung mit der Operationalisierung von Leistung ist demnach geboten, bleibt jedoch häufig aus (Combs et al., 2005). Die hier grundlegende Appropriationstheorie bietet einen empirisch als gehaltvoll bestätigten Zugang zum Verständnis dessen, was mit einer spezifischen Leistungsoperationalisierung überhaupt gemessen wird. Nicht immer ist es ideal, nahe an der Kernwertschöpfung zu messen, je nach Gegenstand des Interesses kann es durchaus sinnvoll sein, Leistungsoperationalisierungen zu wählen, die nur die Appropriationsleistung bestimmter Stakeholder erfassen. Versucht man etwa die Verteilung von Werten in Organisationen oder die Beziehung zwischen einer bestimmten organisationalen Zielfunktion und einer Einflussgröße zu verstehen ist dies sogar der einzig sinnvolle Weg. Ein Streben hin zu einer Art best practice der Leistungsoperationalisierung in Form einer „überlegenen“ Operationalisierung wäre dieser Argumentation folgend nicht nur nicht geboten sondern für den Wissenszuwachs im Fachgebiet sogar schädlich. Aus Sicht der Appropriationstheorie bieten die in dieser Arbeit gewonnenen Erkenntnisse darüber hinaus die Chance, die Wirkung verschiedener erklären-

der Variablen auf die Machtverteilung in Organisationen untersuchen zu können, ein lohnendes Feld für zukünftige Forschungen.

Dennoch ist klar, dass sich bestimmte Operationalisierungen besser eignen um bestimmte Aspekte der Leistung von Organisationen zu erfassen als andere. Verschiedene Autoren (etwa Barney & Mackey, 2005; March & Sutton, 1997) zeigen sich etwa grundsätzlich skeptisch gegenüber der Möglichkeit, „von Außen“ tiefere Erkenntnisse über eine Organisation gewinnen zu können und plädieren daher für einen tieferen Blick in die Organisation hinein, im Falle von Barney & Mackey (2005) in Form von Insider-Befragungen. Die Analyse von Organisationen aus einer reinen Außenperspektive, also durch eine Befragung von der Organisation externen Experten, bietet beispielsweise kaum einen der Vorteile, die interne Experten auszeichnen, und erscheint damit eher suboptimal (vergleiche hierzu auch Lebas & Euske, 2007). Legt man Studien marktbasierter Kennzahlen zu Grunde können die Auswirkungen einer bestimmten erklärenden Variablen möglicherweise mit auf enge Zeitfenster bezogenen Kennzahlen wie CAR besser abgebildet werden als mit solchen, die durch ihren kumulierenden Effekt eine größere Zahl an unerklärten Faktoren einbeziehen (etwa TSR), vorausgesetzt, die erklärende Variable lässt sich auf ein ebenso spezifisches enges Zeitfenster eingrenzen. Ähnliche Überlegungen lassen sich für andere Performanceoperationalisierungen anstellen. Grundsätzlich muss es immer darum gehen, eine Leistungsoperationalisierung zu wählen, die möglichst nur die Zusammenhänge und Aspekte abbildet, die abgebildet werden sollen.

Um die immanenten Probleme jeder Leistungsoperationalisierung zu heilen oder zumindest einzugrenzen bieten sich grundsätzlich zwei Wege an: Zum einen eine neue Bescheidenheit bei Präsentation und insbesondere Zitation von Ergebnissen der Strategischen Managementforschung, die auf lediglich einer oder wenigen Untersuchungen basieren, analog den Überlegungen, die oben zur Entkontextualisierungsdiskussion vorgestellt wurden. Diese Bescheidenheit muss dazu führen, dass Aussagen über kausale Zusammenhänge erst getroffen werden, wenn ein ausreichend großes Sample an Studien zum gleichen Thema, aber mit unterschiedlichen Leistungsoperationalisierungen und Stichproben, vorliegt. Ist dies der Fall können die Ergebnisse in Metaanalysen zusammengeführt werden. Für dieses auch als Triangulation bezeichnete Vorgehen plädieren unter anderem Short et al. (2002), Singh et al. (2003) und Richard et al. (2009).

Auf der Ebene von Einzelstudien ließe sich die Aussagekraft einer empirischen Untersuchung wesentlich erhöhen, wenn die abhängige Variable Leistung auf unterschiedliche Art und Weise operationalisiert und in einen Zusammenhang zu den erklärenden Variablen gestellt werden würde. So kann ein zweites Kriterium geschaffen werden, anhand dessen dann fundierte Aussagen über Zusammenhänge möglich werden. Differenzen in den Ergebnissen könnten dann vor dem Hintergrund eines profunden Verständnisses der jeweiligen Leistungsoperationalisierung theoretisch durchdrungen werden und so die Gesamtaussagekraft des Modells bereichern. Für ein solches mehrgleisiges Vorgehen plädieren beispielsweise auch Boyd et al. (2005) und Takeuchi et al. (2008). Klar ist, dass sich der Aufwand für den Forscher erhöht, gibt es hier doch wenigstens eine Variable mehr, die einer Operationalisierung und insbesondere einer Datensammlung bedarf. Beachtenswert ist, dass dieses Verfahren, wie die Literaturreview in Kapitel 4 zeigte, durchaus bereits Anwendung findet, wenn auch noch nicht in wünschenswerter Häufigkeit und Stringenz. Wichtig ist in diesem Zusammenhang auch die grundsätzliche Anmerkung von Orlitzky et al. (2003), die ebenfalls die Notwendigkeit betonen,

komplexe Konstrukte in unterschiedlicher Art und Weise zu operationalisieren. Nur auf diese Weise lässt sich auch die Multidimensionalität dieser Konstrukte operativ fassen. Dies gilt in besonderer Weise, jedoch nicht ausschließlich, für das Konstrukt Leistung. Ähnlich argumentieren auch Boyd et al. (2005), die betonen, dass sich insbesondere Messfehler durch eine mehrdimensionale Operationalisierung eingrenzen lassen.

In jedem Fall ist ein Verständnis dessen, was eine bestimmte Leistungsoperationalisierung genau misst, von entscheidender Bedeutung für die Interpretation von Forschungsergebnissen. In der vorgelegten Arbeit konnte gezeigt werden, dass die Aufspaltung des Leistungskonstrukts in Leistungserzeugung (Kernperformance) und Leistungsappropriation einen guten Ansatz hierfür liefert.

Zudem muss, zum Ende der Betrachtungen zum Leistungsbegriff, auch ein sozialkonstruktivistisches Argument noch Beachtung finden. Lebas & Euske (2007) betonen den subjektiven Charakter der Bewertung der Leistung einer Organisation. Auch wenn sich diese Argumentation auf die Perspektive des Business Performance Measurement bezieht, hat sie doch auch für die Leistungsmessung aus Sicht der Strategischen Managementforschung Gültigkeit. Die Sinnzuschreibung einer zunächst zumeist allenfalls deskriptiven Kenngröße obliegt dem Beobachter. Dies ist im Rahmen der vergleichenden Leistungsmessung zunächst kein Problem, ist der Beobachter doch (angenommenermaßen) immer derselbe, der Wissenschaftler und sein Team. Ein Problem tritt aber auf, wenn die Leistungskenngröße Bewertungen anderer Personen enthält. Dies ist bereits bei marktbasierten Kennzahlen der Fall, diese bilden die Erwartungen der Marktteilnehmer über die Zukunft des Unternehmens auf der Basis der Interpretation der Betrachtungszeitpunkt verfügbaren Informationen ab (Kabanoff & Brown, 2008). Sie enthalten damit also bereits eine subjektiv-bewertende Komponente, wenn auch über die Zahl der Marktteilnehmer zu einer Art kollektiver Einschätzung ausnivelliert. Dieses Problem verstärkt sich, wenn die Interpretation der Leistung bereits im Zuge der Operationalisierung „externalisiert“ wird, wie es im Rahmen der Befragung von internen und externen Experten der Fall ist. Hier interpretiert der Experte bewusst oder unbewusst ihn erreichende Informationen über das oder aus dem Unternehmen anhand seines spezifischen Deutungsschemas. Er führt einen individuellen Bias in die Operationalisierung ein. Dieser Bias ist bei externen Experten erwartungsgemäß größer als bei internen, da diese die Organisation selbst in mehr oder weniger starkem Grade als Black Box betrachten müssen, während interne Experten auf Erwartungen über Handlungen ihnen bekannter Personen zurückgreifen können, mit denen sie gegebenenfalls auch eine gemeinsame Wirklichkeitskonstruktion teilen (Lebas & Euske, 2007; vergleiche auch Radatz, 2006). Dieses Problem der intersubjektiven Validität soll an dieser Stelle nur angerissen werden und wird bei Nørreklit et al. (2007) und Steigenberger, Pawlowsky, Lübcke und Schulte (2010) ausführlicher diskutiert.

Festzuhalten bleibt, dass die vorliegende Arbeit klare empirische Evidenz für die Relevanz der Messung der organisationalen Leistung in der Strategischen Managementforschung erbracht hat. „Leistung“ besteht aus mindestens zwei getrennt zu betrachtenden Konstrukten, der Schaffung einer Kernperformance sowie deren Verteilung auf unterschiedliche Stakeholderguppen. Die Art der Leistungsmessung erklärt einen wesentlichen Teil der Heterogenität von Forschungsergebnissen in der Strategischen Managementforschung.

6.2.1.3 Epistemologische Implikationen

Fraglich ist, inwieweit sich das in Abschnitt 2.1 erarbeitete Ideal eines Wissensfortschritts durch Falsifikation von Theorien in der Strategischen Managementforschung durchhalten lässt. Wesentlich ist hier die Argumentation von Powell (Powell, 2001; Powell, 2002; Powell, 2003a), die sich teilweise auch in anderen Beiträgen wiederfindet die sich kritisch mit der Strategischen Managementforschung auseinandersetzen (etwa Nicolai & Kieser, 2002; March & Sutton, 1997; auch Gadamer, 1990). Powell lehnt bereits die grundsätzliche Notwendigkeit einer gemeinsamen epistemologischen Basis, wie sie etwa der kritische Rationalismus darstellt, ab und plädiert für einen pragmatischen Empirismus. Powell zeigt auf, dass verschiedene Theoriegebäude in der Strategischen Managementforschung auf völlig unterschiedlichen epistemologischen Fundamenten aufbauen und aus der resultierenden Inkommensurabilität auch kein Schaden erwachsen muss. Damit ist jedoch nicht zu erwarten, dass es jemals gelingen kann, eine Theorie des Wettbewerbsvorteils oder der Hochleistung zu entwickeln. Als Beispiel führt Powell (Powell, 2002, 879) die Unvereinbarkeit des philosophischen Determinismus, der etwa der Populationsökologie, Resource-Dependence-Ansätzen oder marxistischen Theorien und auch dem Market-based view, zugrunde liegt, mit voluntaristischen Ansätzen, die etwa im Bereich der normativen Dimension der Strategischen Managementforschung verständnisleitend sind, an. In letzterem Fall ergibt sich ein bestimmtes Ergebnis des arbeitsteiligen Marktprozesses als Ergebnis der Handlungen im doppelten Wortsinne entscheidender Personen, in ersterem Fall aus Gegebenheiten, die außerhalb des Entscheidungsraumes dieser Personen liegen. Ganz offensichtlich sind die grundlegenden Annahmen der beiden Sichtweisen unvereinbar und können aus einer kritisch-rationalistischen Sicht in der Tradition von Leibniz und Popper nicht gleichzeitig wahr sein. Dennoch lassen sich beide Ansätze empirisch bestätigen, beide leisten einen Beitrag zur Varianzaufklärung in der Strategischen Managementforschung. Powell (2002) folgert überdies aus seinen Überlegungen zum Tautologieproblem des Resource-based view, dass ein Konzept wie Wettbewerbsvorteil sich aufgrund seines unspezifischen Charakters grundsätzlich empirischer Falsifikation verschließt. Er spricht sich daher für einen pragmatischen Ansatz in der Strategischen Managementforschung aus: Da positivistische Ansprüche an die Wahrheit durch Theorien, die auf Wettbewerbsvorteile basieren, nicht erfüllt werden können, sollten Theorien daran gemessen werden, inwieweit sie zur Lösung menschlicher Probleme beitragen.

Die vorliegende Untersuchung bietet für diese Diskussion zunächst wenig grundsätzlichen Neuigkeitswert. Mutig könnte man, im Sinne der Hoffnung von Hunter & Schmidt (2007), annehmen, dass die im Rahmen der Metaanalyse geleistete Varianzaufklärung als Hinweis darauf verstanden werden kann, dass es tatsächlich *eine* Theorie der Leistung von Organisationen geben könnte, die auf kritisch-rationalistischem Wege erschließbar sein könnte. Powells Analyse wäre damit zu widersprechen. Gerade die Multidimensionalität des Leistungskonstrukts, die ja lediglich eine, wenn auch eine zentrale, Variable der Strategischen Managementforschung darstellt, lässt solche Hoffnungen aber kaum zu. In einem System, in dem weder die Konstrukte auf Seiten der abhängigen noch auf Seiten der unabhängigen Variablen eindeutig theoretisch ableitbar, abgrenzbar und operationalisierbar sind und dem darüber hinaus auch eine gemeinsame epistemologische Basis fehlt wird sicherlich weiterhin mit parallelen, untereinander inkommensurablen Theorien und Theoriegruppen gearbeitet und geforscht werden. Im Kuhn'schen Sinne (Kuhn, 2007) liegen also eine ganze Reihe rivalisierender Paradigmen, beziehungsweise vorparadigmatischer Denkschulen, vor, die sich jedoch

kaum gegenseitig aus dem Felde schlagen werden, da gemeinsame Maßstäbe zur Bewertung ihrer Leistungsfähigkeit fehlen. Die Falsifikation von Hypothesen kann hier, aufgrund der genannten Pluralität der Konstrukte, kein Kriterium sein, da durch die Operationalisierung relevanter Konstrukte beinahe alle Theorien in irgend einem Setting falsifiziert wurden, wie in der Arbeit gezeigt werden konnte. Powells pragmatischer Ansatz bleibt also wohl verständnisleitend, wenn auch aus Sicht des kritischen Rationalismus unbefriedigend.

Im Zuge des Paradigmenwettstreits wird auch die Frage spannend bleiben, ob der Ansatz der Komplexitätsreduktion, der die derzeitige Managementforschung dominiert, beherrschend bleibt oder in absehbarer Zukunft an Grenzen stößt. Wie die vorliegende Untersuchung zeigt geht die Reduktion von Komplexität immer mit der Schaffung unbeobachteter Variablen einher, die wiederum die Möglichkeit einschränken, robuste Ergebnisse zu generieren. Die Organisationskybernetik bietet hier einen interessanten alternativen Ansatz, der wiederum jedoch daran krankt, dass er sich quantitativen empirischen Erhebungen eben aufgrund seines Umgangs mit Komplexität bisher weitgehend verschließt. In der jüngsten Zeit wurden einige Beiträge in hochwertigen Zeitschriften publiziert, die als Tendenz hin zu einer eher qualitativen Vorgehensweise verstanden werden können (vergleiche insbesondere Durand & Vaara, 2009). Möglicherweise liegt hierin die Zukunft der empirischen Strategischen Managementforschung.

6.2.2 Grenzen dieser Untersuchung und weiterer Forschungsbedarf

Methodische Limitationen wurden bereits im Rahmen der beiden dieser Arbeit zu Grunde liegenden empirischen Erhebungen in den Abschnitten 4.1.4 und 5.6.1 diskutiert. Generell muss die vorliegende Arbeit naturgemäß die Antwort schuldig bleiben, inwieweit die in einem Einzelfall, der Beziehung zwischen Ressourcen im Sinne des Resource-based view und der Leistung von Organisationen, als relevant aufgezeigten Moderatorvariablen auch in anderen Diskursen im Bereich der Strategischen Managementforschung bedeutsam sind. Folgestudien für weitere, stark beforschte Themenstellungen im Fachgebiet sind notwendig. Allgemein zeigt die Arbeit die Bedeutung der Methodik der Metaanalyse für die Strategische Managementforschung auf.

Ein im Rahmen dieser Arbeit unbeachtet gebliebenes Problem ist die Einbeziehung von Risiko in die Leistungsmessung. Grundsätzlich ist die Leistung einer Organisation höher einzuschätzen, wenn sie einen gegebenen Ertrag mit einem geringeren Risiko erreicht hat als ihre Vergleichsgruppe, da die statistischen Erwartungswerte in diesem Fall höher liegen (vergleiche Hartung et al., 2009), beziehungsweise aus einer ökonomischen Rentabilitätsanalyse heraus der Barwert einer entsprechenden Investition in eine Organisation höher ausfällt.¹³⁶ Dieser Faktor wird im Rahmen der Leistungsoperationalisierung im Fachgebiet bisher gänzlich vernachlässigt (Likierman, 2007) und konnte auch im Rahmen dieser Arbeit nicht erfasst werden. Grundsätzlich müsste die Leistung einer Organisation mit einem Risikodiscount versehen werden, der entsprechende Risiken abbildet. Ein Beispiel wie dies möglich sein könnte ist die Arbeit von Short et al. (2007) zu strategischen Gruppen. Hier wird die Performance der unter-

¹³⁶ Diese Barwertbetrachtung ist vor dem Hintergrund der in Abschnitt 3.3.1 entwickelten Organisationstheorie für jede Person oder Gruppe relevant, die in eine vertragliche Beziehung zur Organisation treten will, also beispielsweise auch Mitarbeiter, Kunden und Lieferanten, vergleiche beispielsweise auch Staehle, Conrad & Sydow, 1999.

suchten Unternehmen explizit mit dem Risiko, das mit der Erzielung der Leistung verbunden war, in Verbindung gesetzt. Die Autoren nutzen hierzu drei Operationalisierungen. Während die accountingbasierte Kennzahl ROA als nicht risikoangepasste Größe gilt, wird Tobins q , also eine marktbasierende Kennzahl, eine gewisse Risikokontrolle zugesprochen. Als maximale Ausprägung der Risikoeinbeziehung nutzen die Autoren schließlich noch Altman's Z , einen Indikator, der die Konkurswahrscheinlichkeit eines Unternehmens messen soll. Marktbasierenden Kennzahlen wird allgemein eine größere Robustheit bei der Messung des inkorporierten Risikos zugeschrieben als accountingbasierten (Kroll et al., 2007; Lee & Makhija, 2009).¹³⁷ Sorenson et al. (2006) plädieren für die Verwendung von „Überleben“ als Leistungsoperationalisierung, da diese im Vergleich zu accountingbasierten Kennzahlen das Risiko deutlich besser abbildet. Die Wahrnehmung von Risiko, die Risikoneigung von Managern und Organisationen und die Beziehung zwischen Risiko und Leistung werden in verwandten Diskursen umfangreich diskutiert.¹³⁸ Eine Einbeziehung in Überlegungen zur Leistungsmessung von Organisationen wäre geboten, hier besteht weiterer Forschungsbedarf. So könnte etwa der aus der Investitionsrechnung entlehnte Realloptionenansatz (etwa Oriani & Sobrero, 2008; Tong, Reuer & Peng, 2008) einen möglichen Zugang zur Abbildung von Risiko in der Strategischen Managementforschung darstellen.

Schließlich ist auch noch anzumerken, dass die Moderatorvariablen, die im Rahmen der vorliegenden Arbeit getestet wurden, zumindest in Teilen nur unzureichend verstanden sind. So sind etwa die entstehenden Time-Lags in Unterbereichen der Strategischen Managementforschung kaum erforscht. Wie lange dauert es etwa, bis sich die Auswechslung eines CEOs oder ein Strategiewechsel in der Performance eines Unternehmens niederschlagen? Dies zu wissen wäre wichtig um auf einer mehr als nur spekulativen Basis für die in dieser Arbeit diskutierte Time-Lag-Problematik zu kontrollieren. Die Appropriationstheorie verschließt sich bisher weitgehend einer direkten empirischen Analyse, da die Messung der Appropriationsleistung von Stakeholdergruppen nach wie vor ein ungelöstes Problem darstellt (vergleiche Blyler & Coff, 2003). Wie im Rahmen dieser Arbeit in Abschnitt 3.3.4 entwickelt würde die Messung der Appropriationsleistung die Erfassung von Minimalpreisen erfordern, zu denen bestimmte Stakeholder gerade noch willens sind, ihre Leistung in die Organisation einzubringen. Ein Indikator, der in der Lage ist, sich diesem Konstrukt anzunähern, wäre notwendig, soll die Appropriationstheorie von einer derzeit weitgehend konzeptionellen Betrachtung in das Stadium einer empirisch überprüfbaren Theorie überführt werden. Hier besteht erheblicher Forschungsbedarf. Tatsächlich ist die vorliegende Arbeit mit ihrem metaanalytischen Ansatz eine der wenigen, die bisher appropriationstheoretische Überlegungen empirisch zu überprüfen in der Lage waren. Von weiteren Forschungen aus dem Verständnis der Appropriationstheorie heraus wären wesentliche Erkenntnisse zu Machtprozessen innerhalb von Unternehmen zu erwarten, hochgradig relevant sowohl aus einer wissenschaftlichen als auch aus einer normativen Perspektive.

¹³⁷ Allerdings ist die Empirie hier nicht eindeutig, so finden beispielsweise Hillman, Shropshire & Canella Jr. (2007) nur einen minimalen Zusammenhang zwischen marktbasierenden Kennzahlen und der Verschuldungsquote von Organisationen, die als ein weiteres Risikomaß verstanden werden kann.

¹³⁸ Vergleiche beispielsweise Andersen et al. (2007) und Henkel (2009) zum Risk-Return-Paradox; Devers, Wiseman & Holmes Jr. (2007), Goranova, Alessandri, Brandes & Dharwadkar (2007) und Simsek (2007) zur Risikoneigung von Managern; Larraza-Kintana, Wiseman, Gomez-Mejia & Welbourne (2007), Wright, Kroll, Krug & Pettus (2007) und Sitkin, Barden, Chattopadhyay, George & Elizabeth, (2006) zur Risikoneigung von Organisationen.

Anhänge

A.1 Verzeichnisse

Abbildungsverzeichnis

Abbildung 1: Ökonomische Rentengenerierung.....	52
Abbildung 2: Stakeholder-Modell	54
Abbildung 3: Flussdiagramm Literaturreview	82
Abbildung 4: Boxplot n.....	116
Abbildung 5: Test auf Normalverteilung der abhängigen Variablen, grafische Auswertung..	118
Abbildung 6: Leistungsoperationalisierungen disaggregiert, deskriptiv.....	130
Abbildung 7: Interdependenz Jahrzehnt / Notierung	138

Tabellenverzeichnis

Tabelle 1: Relative Verhandlungsmacht der internen Stakeholder	55
Tabelle 2: Häufigkeiten Absolutwerte.....	89
Tabelle 3: Häufigkeiten Relativwerte	90
Tabelle 4: Relative Häufigkeiten ohne Nullwert	90
Tabelle 5: Häufigkeit der Operationalisierungen – aggregiert.....	91
Tabelle 6: Häufigkeit der Operationalisierungen – disaggregiert	91
Tabelle 7: Deskriptive Statistik Anzahl der Operationalisierungen.....	92
Tabelle 8: Überblick Selektionsfilter	93
Tabelle 9: Eingrenzung der Aussagemacht diskutiert?	94
Tabelle 10: Häufigkeit der Kontrolle für Time-Lags.....	95
Tabelle 11: Methoden des Umgangs mit Time-Lag Problemen.....	95
Tabelle 12: Optionen für Umgang mit Zeitversatzproblem	95
Tabelle 13: Überblick Methodik Metaanalyse	102
Tabelle 14: Deskriptive Analyse n	115
Tabelle 15: Deskriptive Analyse der abhängigen Variablen „korrigierte Korrelation“	117
Tabelle 16: Kolmogorov-Smirnov-Test auf Normalverteilung	118
Tabelle 17: t-Test, Gruppierungsvariable „Technologieregime“.....	120
Tabelle 18: t-Test, Gruppierungsvariable „Branchen aggregiert“	120
Tabelle 19: Deskriptive Statistik, Gruppierungsvariable „Jahrzehnt“	121
Tabelle 20: t-Test, Gruppierungsvariable „Jahrzehnt“	121
Tabelle 21: Deskriptive Statistik, Gruppierungsvariable „Konjunkturzyklus“	122
Tabelle 22: Levene-Test und ANOVA, Gruppierungsvariable „Konjunkturzyklus“	123
Tabelle 23: Scheffé-Test, Gruppierungsvariable „Konjunkturzyklus“	123

Tabelle 24: Deskriptive Statistik, t-Test und univariate ANOVA, Gruppierungsvariable „Börsennotierung“	124
Tabelle 25: Deskriptive Statistik und Levene-Test, Gruppierungsvariable „Leistungsoperationalisierung aggregiert“	125
Tabelle 26: ANOVA und Scheffé-Test für Gruppierungsvariable „Leistungsoperationalisierung aggregiert“	126
Tabelle 27: Deskriptive Statistik, Gruppierungsvariable „Leistungsoperationalisierungen disaggregiert“	128
Tabelle 28: Levene-Test, Gruppierungsvariable „Leistungsoperationalisierungen disaggregiert“	128
Tabelle 29: ANOVA und Scheffé-Test, Gruppierungsvariable „Leistungsoperationalisierung disaggregiert“	129
Tabelle 30: Deskriptive Statistik, Gruppierungsvariable „Zeitversatz“	131
Tabelle 31: t-Test und Levene-Test, Gruppierungsvariable „Zeitversatz“	132
Tabelle 32: ANOVA, Sample ohne expertenbasierte Leistungsoperationalisierung, Gruppierungsvariable „Zeitversatz“	133
Tabelle 33: Deskriptive Statistik, reduziertes Sample Ressource HPWS, Gruppierungsvariable „Leistungsoperationalisierung aggregiert“	135
Tabelle 34: Levene-Test, ANOVA und Interaktionsmatrix Notierung / Zeitraum	136
Tabelle 35: ANOVA Zeit / Leistungsoperationalisierung	139
Tabelle 36: Interaktionsmatrix Zeit / Time Lag	140
Tabelle 37: Interaktionsmatrix Leistungsoperationalisierung / Time-Lag Robustheit	141
Tabelle 38: Überblick Hypothesentests Metaanalyse	141
Tabelle 39: Rechenbeispiel Appropriation	151
Tabelle 40: Forschungsfragen der qualitativen Literaturreview	162
Tabelle 41: Zusammenfassung Hypothesen Metaanalyse	163

Abkürzungsverzeichnis

AMJ – Academy of Management Journal

ANOVA – Analysis of Variance

CAR – Cumulative Abnormal Returns

CEO – Chief Executive Officer

CSP – Corporate Social Performance

CSR – Corporate Social Responsibility

EVA – Economic Value Added

FuE – Forschung und Entwicklung

HPWS – High Performance Worksystems

HRM – Human Resource Management

IPO – Initial Public Offering

NYSE – New York Stock Exchange

OLS – Ordinary Least Squares, dt: Methode der kleinsten Abweichungsquadrate

PEG – Price Earnings Ratio, dt: Kurs-Gewinn-Verhältnis

ROA – Return on Assets

ROCE – Return on Capital Employed

ROE – Return on Equity

ROI – Return on Investment

ROS – Return on Sales

SMJ – Strategic Management Journal

TSR – Total Shareholder Return; Synonyme: Total Stock Market Return, Total Return to Shareholders. Veränderung der Marktkapitalisierung in einem Jahr plus ausbezahlte (und bei Mehrjahresbetrachtung reinvestierte) Dividenden.

Literaturverzeichnisse

Zur Systematik

Das Literaturverzeichnis ist aus Gründen der Übersichtlichkeit dreigeteilt. Das erste Verzeichnis umfasst Schriften die im Text der Arbeit zitiert wurden, das zweite die Artikel, die in die qualitative Literaturreview einbezogen wurden, das dritte schließlich die Primärstudien der Metaanalyse. Da alle drei Verzeichnisse für sich vollständig sind, sind sie nicht überschneidungsfrei.

Literaturverzeichnis I: In der vorgelegten wissenschaftlichen Arbeit zitierte Quellen

Wissenschaftliche Quellen

- Acquaah, M. (2007). Managerial Social Capital, Strategic Orientation, and Organizational Performance in an Emerging Economy. *Strategic Management Journal*, 28(12), 1235–1255.
- Aggarwal, V. A. & Hsu, D. H. (2009). Modes of Cooperative Commercialization by Start-Ups. *Strategic Management Journal*, 30(8), 835–864.
- Andersen, T. J., Denrell, J. & Bettis, R. A. (2007). Strategic Responsiveness and Bowman's Risk-Return Paradox. *Strategic Management Journal*, 28(4), 407–429.
- Ang, S. Hwee (2008). Competitive Intensity and Collaboration: Impact on Firm Growth Across Technological Environments. *Strategic Management Journal*, 29(10), 1057–1075.
- Arthaud-Day, M. L., Certo, S. Trevis, Dalton, C. M. & Dalton, D. R. (2006). A Changing of the Guard: Executive and Director Turnover Following Corporate Financial Restatements. *Academy of Management Journal*, 49(6), 1119–1136.
- Asher, C. Charleton, Mahoney, J. M. & Mahoney, J. T. (2005). Towards a Property Rights Foundation for a Stakeholder Theory of the Firm. *Journal of Management and Governance*, 9(1), 5–32.
- Atkinson, A. A., Waterhouse, John H. & Wells, Robert B. (1997). A Stakeholder Approach to Strategic Performance Measurement. *Sloan Management Review*, 38(3), 25–37. Retrieved April 29, 2008.
- Bae, J. & Gargiulo, M. (2004). Partner Substitutability, Alliance Network Structure and Firm Profitability in the Telecommunications Industry. *Academy of Management Journal*, 47(6), 843–859.
- Baker, G., Gibbons, R. & Murphy, K. J. (2002). Relational Contracts and the Theory of the Firm. *Quarterly Journal of Economics*, 117(1), 39–84.
- Balkin, D. B., Markman, G. D. & Gomez-Mejia, L. R. (2000). Is CEO Pay in High-Technology Firms Related to Innovation. *Academy of Management Journal*, 43(6), 1118–1129.
- Bamberger, P. (2008). Beyond Contextualization: Using Context Theories to Narrow the Micro-Macro-Gap in Management Research. *Academy of Management Journal*, 51(5), 839–846.
- Barden, J. Q. & Mitchell, W. (2007). Disentangling the Influences of Leaders' Relational Embeddedness on Interorganizational Exchange. *Academy of Management Journal*, 50(6), 1440–1461.
- Barkema, H. G. & Shvyrkov, O. (2007). Does Top Management Team Diversity Promote or Hamper Foreign Expansion? *Strategic Management Journal*, 28(7), 663–680.

- Barnett, M. L. (2007). Stakeholder Influence Capacity and the Variability Of Financial Returns To Corporate Social Responsibility. *The Academy of Management Review (AMR)*, 32(3), 794–816.
- Barnett, M. L. & King, A. A. (2008). Good Fences Make Good Neighbors: A Longitudinal Analysis of an Industry Self-Regulatory Institution. *Academy of Management Journal*, 51(6), 1150–1170.
- Barnett, M. L. & Salomon, R. M. (2006). Beyond Dichotomy: The Curvilinear Relationship Between Social Responsibility and Financial Performance. *Strategic Management Journal*, 27(11), 1101–1122.
- Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17, 99–120.
- Barney, J. B. & Mackey, T. B. (2005). Testing Resource-based Theory. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 2, pp. 1–13). Amsterdam u.a.: Elsevier.
- Barr, P. S. (2004). Current and Potential Importance of Qualitative Methods in Strategy Research. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 165–188). Amsterdam u.a.: Elsevier.
- Bea, F. Xaver, Friedl, B. & Schweitzer, M. (2006). *Allgemeine Betriebswirtschaftslehre: Bd 3: Leistungsprozess* (9., neu bearb. und erw. Aufl., Bd. 3). Stuttgart: Lucius & Lucius.
- Becerra, M. (2008). A Resource-Based Analysis of the Conditions for the Emergence of Profits. *Journal of Management*, 34(6), 1110–1126.
- Becker, F. G. (2003). *Grundlagen betrieblicher Leistungsbeurteilungen: Leistungsverständnis und -prinzip, Beurteilungsproblematik und Verfahrensprobleme* (4., akt. Aufl., Vol. 88). Stuttgart: Schäffer-Poeschel.
- Bergh, D. D., Hanke, R., Balkundi, P., Brown, M. & Chen, X. (2004). An Assessment of Research Designs in Strategic Management Research: The Frequency of Threats to Internal Validity. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 347–363). Amsterdam u.a.: Elsevier.
- Bergh, D. D. & Holbein, G. F. (1997). Assesment and Redirection of Longitudinal Analysis: Demonstration With a Study of the Diversification and Divestiture Relationship. *Strategic Management Journal*, 18(7), 557–571.
- Berrone, P. & Gomez-Mejia, L. R. (2009). Environmental Performance and Executive Compensation: An Integrated Agency-Institutional Perspective. *Academy of Management Journal*, 52(1), 103–126.
- Berry, H. (2006). Shareholder Valuation of Foreign Investment and Expansion. *Strategic Management Journal*, 27(12), 1123–1140.
- Blankart, C. B. (2008). *Öffentliche Finanzen in der Demokratie* (7. vollständig überarbeitete Auflage). München: Vahlen.
- Blyler, M. & Coff, R. W. (2003). Dynamic Capabilities, Social Capital, and Rent Appropriation: Ties That Split Pies. *Strategic Management Journal*, 24(7), 677–686.
- Bosse, D. A., Phillips, R. A. & Harrison, J. S. (2009). Stakeholders, Reciprocity, and Firm Performance. *Strategic Management Journal*, 30(4), 447–456.
- Bourne, L. & Walker, D. H. T. (2005). Visualising and Mapping Stakeholder Influence. *Management Decision*, 43(5), 649–660.
- Bowen, H. P. & Wiersema, M. F. (1999). Matching Method to Paradigm in Strategy Research: Limitations of Cross-Sectional Analysis and Some Methodological Alternatives. *Strategic Management Journal*, 20(7), 625–636.

- Bowen, H. P. & Wiersema, M. F. (2004). Modelling Limited Dependent Variables: Methods and Guidelines for Researchers in Strategy and Management. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 87–134). Amsterdam u.a.: Elsevier.
- Boyd, B. K., Gove, S. & Hitt, M. A. (2005). Construct measurement in strategic management research: illusion or reality? *Strategic Management Journal*, 26(3), 239–257.
- Brammer, S. & Millington, A. (2008). Does it Pay to be Different? An Analysis of the Relationship Between Corporate Social and Financial Performance. *Strategic Management Journal*, 29(12), 1325–1343.
- Breene, T. (2003). Redefining high performance. *Accenture Outlook*, (3), 18–23.
- Breene, T. & Nunes, P. F. (2006). Going the distance: How the world's best companies achieve high performance. *Accenture Outlook*, (3), 8–17.
- Bromiley, P. (2005). *The Behavioral Foundations Of Strategic Management*: Blackwell Publishing.
- Bromiley, P. & Johnson, S. (2005). Mechanisms and Empirical Research. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 2, pp. 15–29). Amsterdam u.a.: Elsevier.
- Brosius, F. (2008). *SPSS 16: Das mitp-Standardwerk*. Heidelberg: mitp.
- Buo, J. Carlos & Satorra, A. (2007). The Persistence of Abnormal Returns at Industry and Firm Levels: Evidence from Spain. *Strategic Management Journal*, 28(7), 707–722.
- Burgelman, R. A. & Grove, A. S. (2007). Let Chaos Reign, Then Rein in Chaos - Repeatedly: Managing Strategic Dynamics for Corporate Longevity. *Strategic Management Journal*, 28(10), 965–979.
- Bushman, B. J. (1994). Vote-Counting Procedures in Meta-Analysis. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 193–213). New York: Russell Sage Foundation.
- Buytendijk, F. (2006). The Five Keys to Building a High-Performance Organization. *Business Performance Management*, (February), 24–30.
- Canella Jr., A. A., Park, J.-H. & Lee, H.-U. (2008). Top Management Team Functional Background Diversity and Firm Performance: Examining the Role of Team Member Colocation and Environmental Uncertainty. *Academy of Management Journal*, 51(4), 768–784.
- Capaldo, A. (2007). Network Structure and Innovation: The Leveraging of a Dual Network as a Distinctive Relational Capability. *Strategic Management Journal*, 28(6), 585–608.
- Capon, N., Farley, J. U. & Hoenig, S. (1990). Determinants of Financial Performance: A Meta-Analysis. *Management Science*, 36(10), 1143–1159.
- Capron, L. & Guillén, M. (2009). National Corporate Governance Institutions and Post-Acquisition Target Reorganization. *Strategic Management Journal*, 30(8), 803–833.
- Carolis, D. Marie de (2003). Competences and Imitability in the Pharmaceutical Industry: An Analysis of Their Relationship with Firm Performance. *Journal of Management*, 29(1), 27–50.
- Carroll, G. R. (1994). A Sociological View on Why Firms Differ. In R. P. Rumelt; D. Schendel & D. J. Teece (Eds.), *Fundamental Issues in Strategy: A Research Agenda* (pp. 271–290). Boston: Harvard Business School Press.
- Castanias, R. P. & Helfat, C. E. (2001). The Managerial Rents Model: Theory and Empirical Analysis. *Journal of Management*, 27(6), 661–678.
- Ceccagnoli, M. (2009). Appropriability, Preemption, and Firm Performance. *Strategic Management Journal*, 30(1), 81–98.

- Certo, S. Trevis & Semadeni, M. (2006). Strategy Research and Panel Data: Evidence and Implications. *Journal of Management*, 32(3), 449–471.
- Chacar, A. S. & Hesterly, W. (2008). Institutional Settings and Rent Appropriation by Knowledge-Based Employees: The Case of the Major League Baseball. *Managerial and Decision Economics*, 29(2-3), 117–136.
- Chakrabarti, A., Singh, K. & Mahmood, I. (2007). Diversification and Performance: Evidence from East Asian Firms. *Strategic Management Journal*, 28(2), 101–120.
- Chakravarthy, B. S. (1986). Measuring strategic performance. *Strategic Management Journal*, 7(5), 437–458.
- Chandler, G. N. & Hanks, S. H. (1994). Market Attractiveness, Resource-Based Capabilities, Venture Strategies, and Venture Performance. *Journal of Business Venturing*, 9(4), 331–349.
- Chang, C.-H., Rosen, C. C. & Levy, P. E. (2009). The Relationship Between Perceptions of Organizational Politics and Employee Attitudes, Strain, and Behaviour: A Meta-Analytical Examination. *Academy of Management Journal*, 52(4), 779–801.
- Chang, S. Jin & Xu, D. (2008). Spillovers and Competition Among Foreign and Local Firms in China. *Strategic Management Journal*, 29(11), 495–518.
- Choi, J. & Wang, H. (2009). Stakeholder Relations and the Persistence of Corporate Financial Performance. *Strategic Management Journal*, 30(8), 895–907.
- Coff, R. W. (1999). When Competitive Advantage Doesn't Lead to Performance: The Resource-Based View and Stakeholder Bargaining Power. *Organization Science*, 10, 119–133.
- Coff, R. W. (2010). The Co-Evolution of Rent Appropriation and Capability Development. *Strategic Management Journal*, 31(7), 711–733.
- Coff, R. W. & Lee, P. M. (2003). Insider Trading as a Vehicle to Appropriate Rent from R&D. *Strategic Management Journal*, 24(2), 183–190.
- Collins, C. J. & Clark, K. D. (2003). Strategic Human Resource Practices, Top Management Team Social Networks, and Firm Performance: The Role of Human Resource Practices in Creating Organizational Competitive Advantage. *Academy of Management Journal*, 46(6), 740–751.
- Collis, D., Young, D. & Goold, M. (2007). The Size, Structure, and Performance of Corporate Headquarters. *Strategic Management Journal*, 28(4), 383–405.
- Combs, J. G., Crook, T. Russell & Shook, C. L. (2005). The Dimensionality of Organizational Performance and its Implications for Strategic Management Research. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 2, pp. 259–286). Amsterdam u.a.: Elsevier.
- Combs, J. G., Liu, Y., Hall, A. & Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59(3), 501–528.
- Connor, T. (2007). Market Orientation and Performance. *Strategic Management Journal*, 28(9), 957–959.
- Cooper, H. & Hedges, L. V. (1994). Research Synthesis as a Scientific Enterprise. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 4–14). New York: Russell Sage Foundation.
- Crook, T. Russell & Combs, J. G. (2007). Sources and Consequences of Bargaining Power in Supply Chains. *Journal of Operations Management*, 25(2), 546–555.
- Crook, T. Russell, Ketchen, D. J., Combs, J. G. & Todd, S. Y. (2008). Strategic Resources and Performance: A Meta-Analysis. *Strategic Management Journal*, 29(11), 1141–1154.

- Datta, D. K., Guthrie, J. P. & Wright, P. M. (2005). Human Resource Management and Labor Productivity: Does Industry Matter? *Academy of Management Journal*, 48(1), 135–145.
- David, P., Bloom, M. & Hillman, A. J. (2007). Investor Activism, Managerial Responsiveness, and Corporate Social Performance. *Strategic Management Journal*, 28(1), 91–100.
- David, P., O'Brien, J. P. & Yoshikawa, T. (2008). The Implications of Dept Heterogeneity for R&D Investment and Firm Performance. *Academy of Management Journal*, 51(1), 165–181.
- Davis, J. H., Schoorman, F. David & Donaldson, L. (1997). Toward a Stewardship Theory of Management. *Academy of Management Review*, 22(1), 20–47.
- Delaney, J. & Huselid, M. A. (1996). The Impact of Human Resource Management Practices on Perception of Organizational Performance. *Academy of Management Journal*, 39(4), 949–969.
- Demsetz, H. & Lehn, K. (1985). The Structure of Corporate Ownership: Causes and Consequences. *Journal of Political Economy*, 93(6), 1155–1177.
- Denzau, A. T. & North, D. C. (1994). Shared Mental Models: Ideologies and Institutions. *Kyklos*, 47(1), 3–31.
- Devers, C. E., Wiseman, R. M. & Holmes Jr., R. Michael (2007). The Effects of Endowment and Loss Aversion in Managerial Stock Option Valuation. *Academy of Management Journal*, 50(1), 191–208.
- Devinney, T. M., Richard, P. J., Yip, G. S. & Johnson, G. (2005). *Measuring organizational performance in management research: a synthesis of measurement challenges and approaches*, from <http://ssrn.com/>.
- Diekmann, A. (1995). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. Reinbeck bei Hamburg: Rowohlt Taschenbuch Verlag.
- DiMaggio, P. & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2), 147–160.
- Donaldson, L. & Davis, J. H. (1991). Stewardship Theory or Agency Theory: CEO Governance and Shareholder Returns. *Australian Journal of Management*, 16(1), 49–65.
- Donaldson, T. & Preston, Lee L. (1995). The Stakeholder Theory of the Corporation: Concepts, Evidence and Implications. *Academy of Management Review*, 20(1), 65–91.
- Doucouliafos, H. & Laroche, P. (2009). Unions and Profits: A Meta-Regression Analysis. *Industrial Relations*, 48(1), 146–184.
- Drinkmann, A. (1990). *Methodenkritische Untersuchungen zur Metaanalyse*. Weinheim: Deutscher Studien Verlag.
- Durand, R. & Vaara, E. (2009). Causation, Counterfactuals, and Competitive Advantage. *Strategic Management Journal*, 30(12), 1245–1264.
- Eccles, R. G. (1995). The Performance Measurement Manifesto. In J. Holloway; J. Lewis & G. Mallory (Eds.), *Performance Measurement and Evaluation* (pp. 5–14). London, Thousand Oaks, New Delhi: Sage.
- Elbanna, S. & Child, J. (2007). Influences on Strategic Decision Effectiveness: Development and Test of an Integrative Model. *Strategic Management Journal*, 28(4), 431–453.
- Ellis, K. M., Reus, T. H. & Lamont, B. T. (2009). The Effects of Procedural and Informational Justice in the Integration of Related Acquisitions. *Strategic Management Journal*, 30(2), 137–161.
- Espejo, R. (1994). What is Systemic Thinking? *System Dynamics Review*, 10(2-3), 199–212.
- Fan, X. & Chen, M. (2001). Parental Involvement and Students' Academic Achievement: A Meta-Analysis. *Educational Psychology Review*, 13(1), 1–22.

- Felin, T. & Foss, N. J. (2005). Strategic Organization: A Field in Search of Micro-Foundations. *Strategic Organization*, 3(4), 441–455.
- Fiss, P. C. & Zajac, E. J. (2006). The Symbolic Management of Strategic Change: Sensegiving via Framing and Decoupling. *Academy of Management Journal*, 49(6), 1173–1193.
- Fitza, M., Matusik, S. F. & Mosakowski, E. (2009). Do VCs Matter? The Importance of Owners on Performance Variance in Start-Up Firms. *Strategic Management Journal*, 30(4), 387–404.
- Flach, J. M. (2008). Mind the Gap: A Skeptical View of Macrocognition. In J. M. Schraagen; L. G. Militello; T. Ormerod & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 27–40). Aldershot England Burlington VT: Ashgate Pub.
- Fontenay, C. C. de & Gans, J. S. (2007). A Bargaining Perspective on Strategic Outsourcing and Supply Competition. *Strategic Management Journal*, 28(8), 819–839.
- Foucault, M. (1970/2003). *Die Ordnung des Diskurses* (Erw. Ausg., 9. Aufl.). Frankfurt am Main: Fischer-Taschenbuch-Verl.
- Franco-Santos, M., Kennerley, M., Micheli, P., Martinez, V., Mason, S. & Marr, B., et al. (2007). Towards a definition of a business performance measurement system. *Industrial Journal of Operations & Production Management*, 27(8), 784–801.
- Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Boston: Pitman.
- Freeman, R. Edward, Wicks, A. C. & Parmar, B. (2004). Stakeholder Theory and "The Corporate Objective Revisited". *Organization Science*, 15(3), 364–369.
- Friedman, M. (1966/1994). The Methodology of Positive Economics. *The Philosophy of Economics: An Anthology*, 2, 180–213.
- Fritsch, M., Brixy, U. & Falck, O. (2006). The Effect of Industry, Region, and Time on New Business Survival: A Multi-Dimensional Analysis. *Review of Industrial Organization*, 28(3), 285–306.
- Fritsch, M., Henning, T., Slavtchev, V. & Steigenberger, N. (2007). *Hochschulen, Innovation, Region: Wissenstransfer im räumlichen Kontext*. Berlin: edition sigma.
- Frooman, J. (1999). Stakeholder Influence Strategies. *Academy of Management Review*, 24(2), 191–205.
- Gadamer, H.-G. (1990). *Wahrheit und Methode: Grundzüge einer philosophischen Hermeneutik* (6.th ed.). Tübingen: Mohr.
- Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological review*, 103(4), 650–669.
- Glaser, B. G & Strauss, A. L (1977). *The discovery of grounded theory: Strategies for qualitative research* ([9. Druck]). New York: Aldine de Gruyter.
- Glunk, U. & Wilderom, C. (1996). Organizational Effectiveness = Corporate Performance?: Why and How Two Research Traditions Need to be Merged. *Research Memorandum*, (715), 1–29.
- Godfrey, P. C., Merrill, C. B. & Hansen, J. M. (2009). The Relationship between Corporate Social Responsibility and Shareholder Value: An Empirical Test of the Risk Management Hypothesis. *Strategic Management Journal*, 30(4), 425–445.
- Goranova, M., Alessandri, T. M., Brandes, P. & Dharwadkar, R. (2007). Managerial Ownership and Corporate Diversification: A Longitudinal View. *Strategic Management Journal*, 28(3), 211–225.
- Goshal, S. (2005). Bad Management Theories are Destroying Good Management Practices. *Academy of Management Learning and Education*, 4(1), 75–91.

- Grahovac, J. & Miller, D. J. (2009). Competitive Advantage and Performance: The Impact of Value Creation and Costliness of Imitation. *Strategic Management Journal*, 30(11), 1192–1212.
- Green, B. N., Johnson, C. D. & Adams, A. (2006). Writing Qualitative Literature Reviews for Peer-Reviewed Journals: Secrets of the Trade. *Journal of Chiropractic Medicine*, 5(3), 101–114.
- Greve, H. R. (2008). Multimarket Contact and Sales Growth: Evidence from Insurance. *Strategic Management Journal*, 29(3), 229–249.
- Greve, H. R. & Goldeng, E. (2004). Longitudinal Analysis in Strategic Management. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 135–163). Amsterdam u.a.: Elsevier.
- Guerrero, S. & Barraud-Didier, V. (2004). High-Involvement Practices and Performance of French Firms. *International Journal of Human Resource Management*, 15(8), 1408–1423.
- Gutenberg, E. (1973). *Grundlagen der Betriebswirtschaftslehre* (24. unveränderte Auflage, Band 1: Die Produktion), Band 1: Die Produktion. Berlin u.a.: Springer.
- Haas, M. R. & Hansen, M. T. (2007). Different Knowledge, Different Benefits: Toward a Productivity Perspective on Knowledge Sharing in Organizations. *Strategic Management Journal*, 28(11), 1133–1153.
- Hagedoorn, J. & Cloudt, M. (2003). Measuring innovative performance: is there an advantage in using multiple indicators? *Research Policy*, 32(8), 1365–1379.
- Hall, J. A. & Tickle-Degnen, L. (1994). Hypotheses and Problems in Research Synthesis. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 17–28). New York: Russell Sage Foundation.
- Hall, P. A & Soskice, D. (2004). An Introduction to Varieties of Capitalism. In P. A. Hall & D. Soskice (Eds.), *Varieties of Capitalism: The institutional foundations of comparative advantage* (pp. 1–68). Oxford: Oxford Univ. Press.
- Harrison, J. S., Bosse, D. A. & Phillips, R. A. (2010). Managing for Stakeholders, Stakeholder Utility Functions, and Competitive Advantage. *Strategic Management Journal*, 31(1), 58–74.
- Hart, O. & Moore, J. (1999). Foundations of Incomplete Contracts. *the Review of Economic Studies*, 66(1), 115–138.
- Hartung, J., Elpelt, B. & Klösener, K.-H. (2009). *Statistik: Lehr- und Handbuch der angewandten Statistik* (15.th ed.). München: Oldenbourg.
- Harvard Business Review (Ed.) (2006). *On the High-Performance Organization*. Boston: Harvard Business School Press.
- Hawawini, G., Subramanian, V. & Verdin, P. (2003). Is Performance Driven by Industry- of Firm-Specific Factors? A New Look at the Evidence. *Strategic Management Journal*, 24(1), 1–16.
- Hayton, J. C. (2005). Competing in the New Economy: The Effect of Intellectual Capital on Corporate Entrepreneurship in High-Technology New Ventures. *R&D Management*, 35(2), 137–155.
- Hedges, L. V. (1994). Statistical Considerations. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 29–38). New York: Russell Sage Foundation.
- Hedges, L. V. & Vevea, J. L. (1998). Fixed- and Random-Effects Models in Meta-Analysis. *Psychological Methods*, 3(4), 486–504.
- Henkel, J. (2009). The Risk-Return Paradox for Strategic Management: Disentangling True and Spurious Effects. *Strategic Management Journal*, 30(3), 287–303.

- Heugens, P. P. M. A. R. & Lander, M. W. (2009). Structure! Agency! (And Other Quarrels): A Meta-Analysis of Institutional Theories of Organizations. *Academy of Management Journal*, 52(1), 61–85.
- Hillman, A. J. & Keim, G. D. (2001). Shareholder Value, Stakeholder Management, and Social Issues: What's the Bottom Line? *Strategic Management Journal*, 22(2), 125–139.
- Hillman, A. J., Shropshire, C. & Canella Jr., A. A. (2007). Organizational Predictors of Woman on Corporate Boards. *Academy of Management Journal*, 50(4), 941–952.
- Hitt, M. A., Boyd, B. K. & Li, D. (2004). The State of Strategic Management Research and a Vision for the Future. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 1–32). Amsterdam u.a.: Elsevier.
- Hmieleski, K. M. & Baron, R. A. (2009). Entrepreneurs' Optimism and New Venture Performance: A Social Cognitive Perspective. *Academy of Management Journal*, 52(3), 473–488.
- Hoetker, G. (2007). The Use of Logit and Probit Models in Strategic Management Research: Critical Issues. *Strategic Management Journal*, 28(4), 331–343.
- Holcomb, T. R., Holmes Jr., R. Michael & Connelly, B. L. (2009). Making the Most of What You Have: Managerial Ability as a Source of Resource Value Creation. *Strategic Management Journal*, 30(5), 457–485.
- Holloway, J., Lewis, J. & Mallory, G. (Eds.) (1995). *Performance Measurement and Evaluation*. London, Thousand Oaks, New Dehli: Sage.
- Hull, C. Eirikur & Rothenberg, S. (2008). Firm Performance: The Interactions of Corporate Social Performance with Innovation and Industry Differentiation. *Strategic Management Journal*, 29(7), 781–789.
- Hult, G. Thomas M. & Ketchen, D. J. (2001). Does Market Orientation Matter? A Test of the Relationship Between Positional Advantage and Performance. *Strategic Management Journal*, 22(9), 899–906.
- Hunter, J. E & Schmidt, F. L (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park: Sage.
- Hunter, J. Edward & Schmidt, F. L (2007). *Methods of meta-analysis: Correcting error and bias in research findings* (2. ed.). Thousand Oaks, Calif.: Sage.
- Huselid, M. A. (1995). The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. *Academy of Management Journal*, 38(3), 635–672.
- Iyer, D. & Miller, K. D. (2008). Performance Feedback, Slack, and the Timing of Acquisitions. *Academy of Management Journal*, 51(4), 808–822.
- Jaros, S. (2008). Labor, Knowledge, and Value in the Workplace: Implications for the Pay of Low-Wage Employees. *Journal of Workplace Rights*, 13(1), 3–16.
- Jensen, M. C., Meckling, W. H., Field, S. & Park, T. C. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4), 305–360.
- Jones, T. M., Felps, W. & Bigley, G. A. (2007). Ethical Theory and Stakeholder-Related Decisions: The Role of Stakeholder Culture. *The Academy of Management Review (AMR)*, 32(1), 137–155.
- Joyce, W. F, Nohria, N. & Roberson, B. (2004). *What really works: The 4+2 formula for sustained business success*. New York, NY: HarperBusiness.
- Kabanoff, B. & Brown, S. (2008). Knowledge Structures of Prospectors, Analyzers, and Defenders: Content, Structure, Stability, and Performance. *Strategic Management Journal*, 29(2), 149–171.

- Kacperczyk, A. (2009). With Greater Power Comes Greater Responsibility? Takeover Protection and Corporate Attention to Stakeholders. *Strategic Management Journal*, 30(3), 261–285.
- Kalyta, P. (2009). Compensation Transparency and Managerial Opportunism: A Study of Supplemental Retirement Plans. *Strategic Management Journal*, 30(4), 405–423.
- Kang, E. (2008). Director Interlocks and Spillover Effects of Reputational Penalties from Financial Reporting Fraud. *Academy of Management Journal*, 51(3), 537–555.
- Kant, I. & Eisler, R. (1930/2009). *Das Kant-Lexikon: Nachschlagewerk zu Kants sämtlichen Schriften, Briefen und handschriftlichem Nachlass*. Retrieved August 05, 2009, from <http://www.textlog.de/rudolf-eisler.html>.
- Karnani, A. (2007). Doing Well by Doing Good - Case Study: "Fair & Lovely" Whitening Cream. *Strategic Management Journal*, 28(13), 1351–1357.
- Katzenbach, J. R. (2000). *Peak performance: Aligning the hearts and minds of your employees*. Boston, Mass.: Harvard Business School Press.
- Keil, T., Maula, M., Schildt, H. & Zhara, S. A. (2007). The Effect of Governance Modes and Relatedness of External Business Development Activities on Innovative Performance. *Strategic Management Journal*, 28(8), 895–907.
- Ketchen, D. J. & Bergh, D. D. (Eds.) (2004). *Research Methodology in Strategy and Management*, 1. Amsterdam u.a.: Elsevier.
- Ketchen, D. J. & Bergh, D. D. (Eds.) (2005). *Research Methodology in Strategy and Management*, 2. Amsterdam u.a.: Elsevier.
- Ketchen, D. J., Boyd, B. K. & Bergh, D. D. (2008). Research Methodology in Strategic Management: Past Accomplishments and Future Challenges. *Organizational Research Methods*, 11(4), 643–658.
- Ketchen, D. J., Hult, G. Thomas M. & Slater, S. F. (2007). Toward Greater Understanding of Market Orientation and the Resource-Based View. *Strategic Management Journal*, 28(9), 961–964.
- Kim, J.-Y. (Jay) & Finkelstein, S. (2009). The Effects of Strategic and Market Complementarity on Acquisition Performance: Evidence from the U.S. Commercial Banking Industry, 1989–2001. *Strategic Management Journal*, 30(6), 617–646.
- King, B. (2008). A Social Movement Perspective of Stakeholder Collective Action and Influence. *Business & Society*, 47(1), 21–49.
- Kitchin, J. (1923). Cycles and Trends in Economic Factors. *The Review of Economics and Statistics*, 5(1), 10–16.
- Klein, G. A. (1995). A Recognition-Primed Decision (RPD) Model of Rapid Decision Making. In G. A. Klein; J. Orasanu; R. Calderwood & C. E. Zsombok (Eds.), *Decision Making in Action: Models and Methods* (2nd ed., pp. 138–148). Norwood: Ablex Publishing.
- Kochan, T. A. & Rubinstein, S. A. (2000). Toward a Stakeholder Theory of the Firm: The Saturn Partnership. *Organization Science*, 11(4), 367–386.
- Koka, B. R. & Prescott, J. E. (2008). Designing alliance networks: the influence of network position, environmental change, and strategy on firm performance. *Strategic Management Journal*, 29(6), 639–661.
- Krishnan, R., Martin, X. & Noorderhaven, N. G. (2006). When Does Trust Matter to Alliance Performance? *Academy of Management Journal*, 49(5), 894–917.
- Kroll, M., Walters, B. A. & Le, S. A. (2007). The Impact of Board Composition and Top Management Team Ownership Structure on Post-IPO Performance in Young Entrepreneurial Firms. *Academy of Management Journal*, 50(5), 1198–1216.

- Kronborg, D. & Thomsen, S. (2009). Foreign Ownership and Long-Term Survival. *Strategic Management Journal*, 30(2), 207–219.
- Kuhn, T. S. (2007). *Die Struktur wissenschaftlicher Revolutionen* (2., rev. und um das Postskriptum von 1969 erg. Aufl., [Nachdr.], Vol. 25). Frankfurt am Main: Suhrkamp.
- Lamberg, J.-A., Tikkanen, H., Nokelainen, T. & Suur-Inkeroinen, H. (2009). Competitive Dynamics, Strategic Consistency, and Organizational Survival. *Strategic Management Journal*, 30(1), 45–60.
- Lamnek, S. (1995). *Qualitative Sozialforschung: Methodologie* (3.th ed., Vol. 1), 1. Weinheim: Psychologie Verlags Union.
- Larraza-Kintana, M., Wiseman, R. M., Gomez-Mejia, L. R. & Welbourne, T. M. (2007). Disentangling Compensation and Employment Risks Using the Behavioral Agency Model. *Strategic Management Journal*, 28(10), 1001–1019.
- Lavie, D. (2007). Alliance Portfolios and Firm Performance: A Study of Value Creation and Appropriation in the U.S. Software Industry. *Strategic Management Journal*, 28(12), 1187–1212.
- Lebas, M. J. (1995). Performance Measurement and Performance Management. *International Journal of Production Economics*, 41, 23–35.
- Lebas, M. J. & Euske, K. (2007). A conceptual and operational delineation of performance. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed., pp. 125–142). Cambridge: Cambridge University Press.
- Lee, S.-H. & Makhija, M. (2009). Flexibility in Internationalization: Is it Valuable During an Economic Crisis? *Strategic Management Journal*, 30(5), 537–555.
- Lehmann-Waffenschmidt, M. (2006). Konstruktivismus und Evolutorische Ökonomik. In G. Rusch (Ed.), *Konstruktivistische Ökonomik* (pp. 27–54). Marburg: Metropolis.
- Leiblein, M. J. & Madsen, T. L. (2009). Unbundling Competitive Heterogeneity: Incentive Structures and Capability Influences on Technological Innovation. *Strategic Management Journal*, 30(7), 711–735.
- Leibniz, G. Wilhelm (1704/1961). *Neue Abhandlungen über den menschlichen Verstand*. Frankfurt am Main: Insel Verlag.
- Lepak, D. P., Smith, K. G. & Taylor, M. Susan (2007). Value Creation and Value Capture: A Multilevel Perspective. *Academy of Management Review*, 32(1), 180–194.
- Levitas, E. & McFayden, M. Ann (2009). Managing Liquidity in Research-Intensive Firms: Signaling and Cash-Flow Effects of Patents and Alliance Activities. *Strategic Management Journal*, 30(6), 659–678.
- Li, J., Lam, K. & Qian, G. (2001). Does Culture Affect Behavior and Performance of Firms? The Case of Joint Ventures in China. *Journal of International Business Studies*, 32(1), 115–131.
- Likierman, A. (2007). Risk in performance measurement. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed., pp. 261–277). Cambridge: Cambridge University Press.
- Ling, Y., Simsek, Z., Lubatkin, M. & Veiga, J. F. (2008). Transformational Leadership's Role in Promoting Corporate Entrepreneurship: Examining the CEO-TMT Interface. *Academy of Management Journal*, 51(3), 557–576.
- Lippman, S. A. & Rumelt, R. P. (2003). A Bargaining Perspective on Resource Advantage. *Strategic Management Journal*, 24(11), 1069–1086.
- Lipsey, M. W. (1994). Identifying Potentially Interesting Variables and Analysis Opportunities. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 111–123). New York: Russell Sage Foundation.

- Lorenz, E. (1999). Trust, Contract, and Economic Cooperation. *Cambridge Journal of Economics*, 23, 301–315.
- Lounsbury, M. (2007). A Tale of Two Cities: Competing Logics and Practice Variation in the Professionalizing of Mutual Funds. *Academy of Management Journal*, 50(2), 289–307.
- Luo, Y. (2007a). Are Joint Venture Partners More Opportunistic in a More Volatile Environment? *Strategic Management Journal*, 28(1), 39–60.
- Luo, Y. (2007b). The Independent and Interactive Roles of Procedural, Distributive, and Interactional Justice in Strategic Alliances. *Academy of Management Journal*, 50(3), 644–664.
- Luo, Y. (2008a). Procedural Fairness and Interfirm Cooperation in Strategic Alliances. *Strategic Management Journal*, 29(1), 27–46.
- Luo, Y. (2008b). Structuring Interorganizational Cooperation: The Role of Economic Integration in Strategic Alliances. *Strategic Management Journal*, 29(6), 617–637.
- Lyons, B. & Mehta, J. (1997). Contracts, Opportunism and Trust: Self Interest and Social Orientation. *Cambridge Journal of Economics*, 21, 239–257.
- Macher, J. T. & Boerner, C. S. (2006). Experience and Scale and Scope Economies: Trade-Offs and Performance in Development. *Strategic Management Journal*, 27(9), 845–865.
- Mackey, A. (2008). The Effect of CEOs on Firm Performance. *Strategic Management Journal*, 29(12), 1357–1367.
- Makri, M., Lane, P. J. & Gomez-Mejia, L. R. (2006). CEO Incentives, Innovation, and Performance in Technology-Intensive Firms: A Reconciliation of Outcome and Behavior-Based Incentive Schemes. *Strategic Management Journal*, 27(11), 1057–1080.
- Marcel, J. J. (2009). Why Top Management Team Characteristics Matter When Employing a Chief Operating Officer: A Strategic Contingency Perspective. *Strategic Management Journal*, 30(6), 647–658.
- March, J. G. & Sutton, R. I. (1997). Organizational performance as a dependent variable. *Organization Science*, 8(6), 698–706.
- Marginson, D. & McAulay, L. (2008). Exploring the Debate on Short-Termism: A Theoretical and Empirical Analysis. *Strategic Management Journal*, 29(4), 273–292.
- Martin, K. D., Cullen, J. B., Johnson, J. L. & Parboteeah, K. Praveen (2007). Deciding to Bribe: A Cross-Level Analysis of Firm and Home Country Influences on Bribery Activity. *Academy of Management Journal*, 50(6), 1401–1422.
- Maurer, R. Willi (2004). *Zwischen Erkenntnisinteresse und Handlungsbedarf: Eine Einführung in die methodologischen Probleme der Wirtschaftswissenschaft*. Marburg: Metropolis-Verl.
- McDonald, M. L., Khanna, P. & Westphal, J. D. (2008). Getting Them to Think Outside the Circle: Corporate Governance, CEOs' External Advice Networks, and Firm Performance. *Academy of Management Journal*, 51(3), 453–475.
- McDonald, M. L., Westphal, J. D. & Graebner, M. E. (2008). What Do They Know? The Effects of Outside Director Acquisition Experience on Firm Acquisition Performance. *Strategic Management Journal*, 29(11), 1155–1177.
- McNamara, G. M., Halebian, J. (John) & Dykes, B. Johnson (2008). The Performance Implications of Participating in an Acquisition Wave: Early Mover Advantages, Bandwagon Effects, and the Moderating Influence of Industry Characteristics and Acquirer Tactics. *Academy of Management Journal*, 51(1), 113–130.
- Menguc, B. & Barker, T. (2005). Re-examining Field Sales Unit Performance: Insights From the Resource-Based View and Dynamic Capabilities Perspectives. *European Journal of Marketing*, 39(7/8), 885–909.

- Mesquita, L. F., Anand, J. & Brush, T. H. (2008). Comparing the Resource-based and Relational View: Knowledge Transfer and Spillover in Vertical Alliances. *Strategic Management Journal*, 29(9), 913–941.
- Mesquita, L. F. & Lazzarini, S. G. (2008). Horizontal and Vertical Relationships in Developing Economies: Implications for SMEs' Access to Global Markets. *Academy of Management Journal*, 51(2), 359–380.
- Meyer, M. (2007). Finding performance: the new discipline in management. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed.). Cambridge: Cambridge University Press.
- Michalisin, M. D., Karau, S. J. & Tangpong, C. (2004). Top Management Team Cohesion and Superior Industry Returns: An Empirical Study of the Resource-Based View. *Group and Organization Management*, 29(1), 125–140.
- Miller, D. & Shamsie, J. (1996). The Resource-Based View of the Firm in Two Environments: The Hollywood Film Studios from 1936 to 1965. *Academy of Management Journal*, 39(3), 519–543.
- Miller, D. J., Fern, M. J. & Cardinal, L. (2007). The Use of Knowledge for Technological Innovation Within Diversified Firms. *Academy of Management Journal*, 50(2), 308–326.
- Miller, K. D. (2005). The Problem of Method and the Practice of Management Research. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 2, pp. 143–177). Amsterdam u.a.: Elsevier.
- Mitchell, R. K., Agle, B. R. & Wood, D. J. (1997). Toward a Theory of Stakeholder Identification and Salience: Defining the Principle of Who and What Really Counts. *Academy of Management Review*, 22(4), 853–886.
- Morgan, N. A., Vorhies, D. W. & Mason, C. H. (2009). Market Orientation, Marketing Capabilities, and Firm Performance. *Strategic Management Journal*, 30(8), 909–920.
- Morrow Jr., J. L., Sirmon, D. G., Hitt, M. A. & Holcomb, T. R. (2007). Creating Value in the Face of Declining Performance: Firm Strategies and Organizational Recovery. *Strategic Management Journal*, 28(3), 271–283.
- Mosier, K. (2008). Technology and "Naturalistic" Decision Making: Myths and Realities. In J. M. Schraagen; L. G. Militello; T. Ormerod & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 41–54). Aldershot England Burlington VT: Ashgate Pub.
- Murillo-Luna, J. L., Garcés-Ayerbe, C. & Rivera-Torres, P. (2008). Why Do Patterns of Environmental Response Differ? A Stakeholders' Pressure Approach. *Strategic Management Journal*, 29(11), 1225–1240.
- Nadkarni, S. & Barr, P. S. (2008). Environmental Context, Managerial Cognition, and Strategic Action: An Integrated View. *Strategic Management Journal*, 29(13), 1395–1427.
- Nag, R., Hambrick, D. C. & Chen, M.-J. (2007). What Is Strategic Management, Really? Inductive Derivation of a Consensus Definition of the Field. *Strategic Management Journal*, 28(9), 935–955.
- Neely, A., Kennerley, M. & Adams, C. (2007). Performance measurement frameworks: a review. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed., pp. 143–162). Cambridge: Cambridge University Press.
- Nerur, S. P., Rasheed, A. A. & Natarajan, V. (2008). The Intellectual Structure of the Strategic Management Field: An Author Co-Citation Analysis. *Strategic Management Journal*, 29(3), 319–336.
- Newbert, S. L. (2007). Empirical Research on the Resource-Based View of the Firm: An Assessment and Suggestions for Future Research. *Strategic Management Journal*, 28(2), 121–146.

- Newbert, S. L. (2008). Value, Rareness, Competitive Advantage, and Performance: A Conceptual-Level Empirical Investigation of the Resource-based View of the Firm. *Strategic Management Journal*, 29(7), 745–768.
- Ng, D., Westgren, R. & Sonka, S. (2009). Competitive Blind Spots in an Institutional Field. *Strategic Management Journal*, 30(4), 349–369.
- Ngah-Kiing Lim, E., Das, S. S. & Das, A. (2009). Diversification Strategy, Capital Structure, and the Asian Financial Crisis (1997-1998): Evidence from Singapore Firms. *Strategic Management Journal*, 30(6), 577–594.
- Nicolai, A. & Kieser, A. (2002). Trotz eklatanter Erfolglosigkeit: Die Erfolgsfaktorenforschung weiter auf Erfolgskurs. *Die Betriebswirtschaft (DBW)*, 62(6), 579–596.
- Nordenflycht, A. von (2007). Is Public Ownership Bad for Professional Service Firms? Ad Agency Ownership, Performance, and Creativity. *Academy of Management Journal*, 50(2), 429–445.
- Nørreklit, H., Nørreklit, L. & Mitchell, F. (2007). Theoretical conditions for validity in accounting performance measurement. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed., pp. 179–217). Cambridge: Cambridge University Press.
- Opp, K.-D. (1970). *Methodologie der Sozialwissenschaften*. Reinbeck bei Hamburg: Rowohlt Taschenbuch Verlag.
- Oriani, R. & Sobrero, M. (2008). Uncertainty and the Market Valuation of R&D Within a Real Options Logic. *Strategic Management Journal*, 29(4), 343–361.
- Orlitzky, M., Schmidt, F. L. & Rynes, S. L. (2003). Corporate Social and Financial Performance: A Meta-Analysis. *Organization Studies*, 24(3), 403–441.
- Orwin, R. G. (1994). Evaluating Coding Decisions. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 139–162). New York: Russell Sage Foundation.
- O'Shaughnessy, K. C., Gedajlovic, E. & Reinmoeller, P. (2007). The Influence of Firm, Industry and Network on the Corporate Social Performance of Japanese Firms. *Asia Pacific Journal of Management*, 24(3), 283–303.
- Ozcan, P. & Eisenhardt, K. M. (2009). Origin of Alliance Portfolios: Entrepreneurs, Network Strategies, and Firm Performance. *Academy of Management Journal*, 52(2), 246–279.
- Pacheco-de-Almeida, G., Henderson, J. E. & Cool, K. O. (2008). Resolving the Commitment versus Flexibility Trade-Off: The Role of Resource Accumulation Lags. *Academy of Management Journal*, 51(3), 517–536.
- Pawłowsky, P., Reinhardt, R., Bornemann, M. & Schneider, U. (2003). Intellectual Capital and Knowledge Management: Perspectives on Measuring Knowledge. In M. Dierkes; A. B. Antal; J. Child & I. Nonaka (Eds.), *Handbook of organizational learning and knowledge* (pp. 794–820). Oxford: Oxford Univ. Press.
- Pawłowsky, P. (2008). Auf dem Weg zu höherer Leistung... In P. Pawłowsky & P. Mistele (Eds.), *Hochleistungsmanagement: Möglichkeiten zur Nutzung vorhandener Leistungspotenziale in Organisationen*. Wiesbaden: Gabler.
- Pawłowsky, P. & Mistele, P. (2008). Über den Tellerrand schauen. In P. Pawłowsky & P. Mistele (Eds.), *Hochleistungsmanagement: Möglichkeiten zur Nutzung vorhandener Leistungspotenziale in Organisationen* (pp. 1–18). Wiesbaden: Gabler.
- Pawłowsky, P. & Wilkens, U. (1997). Human Resource Management or Machines That Change the World in the Automotive Industry? *Management International Review*, 37(1), 105–126.

- Payne, A., Holt, S. & Frow, P. (2000). Integrating Employee, Customer and Shareholder Value Through an Enterprise Performance Model: An Opportunity for Financial Services. *International Journal of Bank Marketing*, 18(6), 258–273.
- Pearl, J. (2001). *Causality: Models Reasoning, and inference*. Cambridge: Cambridge Univ Press.
- Peng, M. W. & York, A. S. (2001). Behind Intermediary Performance in Export Trade: Transactions, Agents, and Resources. *Journal of International Business Studies*, 32(2), 327–346.
- Perry-Smith, J. E. & Blum, T. C. (2000). Work-Family Human Resource Bundles and Perceived Organizational Performance. *Academy of Management Journal*, 43(6), 1107–1117.
- Peteraf, M. A. (1993). The Cornerstones of Competitive Advantage: A Resource-based view. *Strategic Management Journal*, 14, 179–191.
- Peteraf, M. A. & Barney, J. B. (2003). Unraveling the Resource-Based Tangle. *Managerial and Decision Economics*, 24(4), 309–323.
- Peters, T. J. & Waterman, R. H. (1982). *In Search of Excellence: Lessons from America's Best-Run Companies*. New York: Harper & Row.
- Pfeffer, J. (1994). *Managing with power: Politics and influence in organizations*. Harvard: Harvard Business School Press.
- Picot, A., Dietl, H. & Franck, E. (2008). *Organisation: Eine ökonomische Perspektive* (5., aktualisierte und überarb. Aufl.). Stuttgart: Schäffer-Poeschel.
- Pike, S. & Roos, G. (2007). The validity of measurement frameworks: measurement theory. In A. Neely (Ed.), *Business Performance Measurement: Unifying theories and integrating practice* (2nd ed., pp. 218–235). Cambridge: Cambridge University Press.
- Plender, J. (1998). Giving People a Stake in the Future. *Long Range Planning*, 31(2), 211–217.
- Popper, K. R. (1973). *Objektive Erkenntnis: Ein evolutionärer Entwurf*. Hamburg: Hoffmann und Campe.
- Popper, K. R. (1989). *Logik der Forschung* (9. verbesserte Auflage). Tübingen: Mohr.
- Porter, M. E. (1999). *Wettbewerbsstrategie: Methoden zur Analyse von Branchen und Konkurrenten* (10., durchges. und erw. Aufl.). Frankfurt: Campus-Verl.
- Powell, T. C. (1992a). Organizational Alignment as Competitive Advantage. *Strategic Management Journal*, 13(2), 119–134.
- Powell, T. C. (1992b). Strategic Planning as Competitive Advantage. *Strategic Management Journal*, 13(7), 551–558.
- Powell, T. C. (2001). Competitive advantage: Logical and philosophical considerations. *Strategic Management Journal*, 22(9), 875–888.
- Powell, T. C. (2002). The Philosophy of Strategy. *Strategic Management Journal*, 23(9), 873–880.
- Powell, T. C. (2003a). Strategy Without Ontology. *Strategic Management Journal*, 24(3), 285–291.
- Powell, T. C. (2003b). Varieties of Competitive Parity. *Strategic Management Journal*, 24(1), 61–86.
- Powell, T. C. & Arregle, J.-L. (2007). Firm Performance and the Axis of Error. *Journal of Management Research*, 7(2), 59–77.
- Pruckner, M. (2002). *Ich wusste gar nicht, dass ich Prosa sprechen kann! Ein Essay über die Management-Kybernetik*.
- Radatz, S. (2006). Re-Culturing. In G. Rusch (Ed.), *Konstruktivistische Ökonomik* (pp. 159–179). Marburg: Metropolis.

- Ray, G., Barney, J. B. & Muhanna, W. A. (2004). Capabilities, business processes, and competitive advantage: choosing the dependent variable in empirical tests of the resource-based view. *Strategic Management Journal*, 25(1), 23–37.
- Reid, J. & Hubbel, V. (2005). Creating a performance culture. *Ivey Business Journal*, (March/April), 1–7.
- Richard, O. C. (2000). Racial Diversity, Business Strategy, and Firm Performance: A Resource-Based View. *Academy of Management Journal*, 43(2), 164–177.
- Richard, O. C., Murthi, B. P. S. & Ismail, K. (2007). The Impact of Racial Diversity on Intermediate and Long-Term Performance: The Moderating Role of Environmental Context. *Strategic Management Journal*, 28(12), 1213–1233.
- Richard, P. J., Devinney, T. M., Yip, G. S. & Johnson, G. (2009). Measuring Organizational Performance: Towards Methodological Best Practice. *Journal of Management*, 35(3), 718–804.
- Robins, J. A. (2004). When Does the Age of Data Matter? Notes on the Selection of Data for Strategy Research. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 251–271). Amsterdam u.a.: Elsevier.
- Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research*. Newbury Park, London, New Delhi: Sage.
- Rowe, W. Glen & Morrow Jr., J. L. (1999). A Note on the Dimensionality of the Firm Financial Performance Construct Using Accounting, Market, and Subjective Measures. *Canadian Journal of Administrative Science*, 16(1), 58–70.
- Rowley, T. J. (1997). Moving Beyond Dyadic Ties: A Network Theory of Stakeholder Influence. *Academy of Management Review*, 22(4), 887–910.
- Rumelt, R. P., Schendel, D. & Teece, D. J. (1991). Strategic Management and Economics. *Strategic Management Journal*, 12, 5–29.
- Rumelt, R. P., Schendel, D. & Teece, D. J. (1994). Fundamental Issues in Strategy. In R. P. Rumelt; D. Schendel & D. J. Teece (Eds.), *Fundamental Issues in Strategy: A Research Agenda* (pp. 9–53). Boston: Harvard Business School Press.
- Rynes, S. L., Giluk, T. L. & Brown, K. G. (2007). The Very Separate Worlds of Academic and Practitioner Periodicals in Human Resource Management: Implications for Evidence-Based Management. *Academy of Management Journal*, 50(5), 987–1008.
- Sampson, R. C. (2004). The Cost of Misaligned Governance in R&D-Alliances. *Journal of Law, Economics and Organization*, 20(2), 484–526.
- Sanders, W. Gerard & Hambrick, D. C. (2007). Swinging for the Fences: The Effects of CEO Stock Options on Company Risk Taking and Performance. *Academy of Management Journal*, 50(5), 1055–1078.
- Schilling, M. A. (2009). Understanding the Alliance Data. *Strategic Management Journal*, 30(3), 233–260.
- Schmalen, H. (2002). *Grundlagen und Probleme der Betriebswirtschaft* (12th ed.). Stuttgart: Schäffer-Poeschel.
- Schnell, R. (1991). Wer ist das Volk? Zur faktischen Grundgesamtheit bei "allgemeinen Bevölkerungsumfragen": Undercoverage, Schwererreichbare und Nichtbefragbare. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 43(1), 106–137.
- Schöler, K. (2004). *Grundlagen der Mikroökonomik: Eine Einführung in die Theorie der Haushalte, der Firmen und des Marktes* (2., überarb. und erw. Aufl.). München: Vahlen.
- Schomann, M. (2001). *Wissensorientiertes Performance Measurement*. Wiesbaden: Deutscher Universitätsverlag.

- Schraagen, J. M., Militello, L. G., Ormerod, T. & Lipshitz, R. (Eds.) (2008). *Naturalistic decision making and macrocognition*. Aldershot England Burlington VT: Ashgate Pub.
- Schreyögg, G. & Kliesch-Eberl, M. (2007). How Dynamic Can Organizational Capabilities Be? Towards a Dual Process Model of Capability Dynamization. *Strategic Management Journal*, 28(9), 913–933.
- Schulze, G., Warning, S. & Wiermann, C. (2008). Zeitschriftenrankings für die Wirtschaftswissenschaften - Konstruktion eines umfassenden Metaindexes. *Perspektiven der Wirtschaftspolitik*, 9(3), 286–305.
- Schumann, O. J. (2006). Wirtschaftsethische Aspekte ökonomischer Theoriebildung. In G. Rusch (Ed.), *Konstruktivistische Ökonomik* (pp. 55–91). Marburg: Metropolis.
- Schumpeter, J. (1952). *Theorie der wirtschaftlichen Entwicklung* (5.th ed.). Berlin: August Raabe.
- Schwaninger, M. (2004). Methodologies in Conflict: Achieving Synergies Between System Dynamics and Organizational Cybernetics. *Systems Research and Behavioral Science*, 21, 411–431.
- Scollon, R. & Wong Scollon, S. (2008). *Intercultural communication: A discourse approach* (2. ed.,). Oxford: Blackwell.
- Seisreiner, A. (2006). *Rationalität wertorientierter Managementkonzepte*. Universität Potsdam, Habilitationsschrift.
- Shanteau, J. (1992). How Much Information Does an Expert Use? Is it Relevant? *Acta Psychologica*, 81(1), 75–86.
- Sharfman, M. P. & Fernando, C. S. (2008). Environmental Risk Management and the Cost of Capital. *Strategic Management Journal*, 29(6), 569–592.
- Shiller, R. J. (2003). From Efficient Markets Theory to Behavioral Finance. *Journal of Economic Perspectives*, 17(1), 83–104.
- Shimizu, K. (2007). Prospect Theory, Behavioral Theory, and the Threat-Rigidity- Thesis: Combinative Effects on Organizational Decisions to Divest Formerly Acquired Units. *Academy of Management Journal*, 50(6), 1495–1514.
- Short, J., Ketchen, D. J. & Palmer, T. B. (2002). The Role of Sampling in Strategic Management Research on Performance: A Two-Study Analysis. *Journal of Management*, 28(3), 363–385.
- Short, J., Ketchen, D. J., Palmer, T. B. & Hult, G. Thomas M. (2007). Firm, Strategic Group, and Industry Influences on Performance. *Strategic Management Journal*, 28(2), 147–167.
- Short, J., Palmer, T. B. & Ketchen, D. J. (2002). Resource-Based and Strategic Group Influences on Hospital Performance. *Health Care Management Review*, 27(4), 7–17.
- Siebert, W. Stanley & Zubanov, N. (2009). Searching for the Optimal Level of Employee Turnover: A Study of a Large U.K. Retail Organization. *Academy of Management Journal*, 52(2), 294–313.
- Simon, H. A. (1946). The Proverbs of Administration. *Public Administration Review*, 6(1), 53–67.
- Simon, H. A. (1952). A Comparison of Organization Theories. *the Review of Economic Studies*, 20(1), 40–48.
- Singh, K., Ang, S. Hwee & Leong, S. Meng (2003). Increasing Replication for Knowledge Accumulation in Strategy Research. *Journal of Management*, 29(4), 533–549.
- Sinha, R. K. & Noble, C. H. (2008). The Adoption of Radical Manufacturing Technologies and Firm Survival. *Strategic Management Journal*, 29(9), 943–962.
- Sirmon, D. G., Gove, S. & Hitt, M. A. (2008). Resource Management in Dyadic Competitive Rivalry: The Effects of Resource Bundling and Deployment. *Academy of Management Journal*, 51(5), 919–935.

- Sitkin, S. B., Barden, J., Chattopadhyay, P., George & Elizabeth (2006). Cognitive Underpinnings of Institutional Persistence and Change: A Framing Perspective. *Academy of Management Review*, 31(2), 347–365.
- Skaggs, B. C. & Youndt, M. (2004). Strategic Positioning, Human Capital, and Performance in Service Organisations: A Customer Interaction Approach. *Strategic Management Journal*, 25(1), 85–99.
- Skilton, P. F. (2009). Knowledge Based Resources, Property Based Resources and Supplier Bargaining Power in Hollywood Motion Picture Projects. *Journal of Business Research*, 62(8), 834–840.
- Slater, S. F. & Athuahene-Gima, K. (2004). Conducting Survey Research in Strategic Management. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 227–249). Amsterdam u.a.: Elsevier.
- Somaya, D., Williamson, I. O. & Lorinkova, N. (2008). Gone but not Lost: The Different Performance Impacts of Employee Mobility Between Cooperators versus Competitors. *Academy of Management Journal*, 51(5), 936–953.
- Sorenson, O., McEvily, S., Rongrong Ren, C. & Roy, R. (2006). Niche Width Revisited: Organizational Scope, Behavior and Performance. *Strategic Management Journal*, 27(10), 915–936.
- Stahle, W. H., Conrad, P. & Sydow, J. (1999). *Management* (8. Aufl. /). München: Vahlen.
- Stam, W. & Elfring, T. (2008). Entrepreneurial Orientation and New Venture Performance: The Moderating Role of Intra- and Extraindustry Social Capital. *Academy of Management Journal*, 51(1), 91–111.
- Starbuck, W. H. (2004). Methodological Challenges Posed by Measures of Performance. *Journal of Management and Governance*, 8(4), 337–343.
- Steigenberger, N., Pawlowsky, P., Lübcke, T. & Schulte, M. (2010). Leistungsmessung in der Strategischen Managementforschung – Methoden der Selektion von Hochleistern. In B. Schallock (Ed.), *Tagungsband des Förderschwerpunkts "Innovationsstrategien jenseits traditionellen Managements" (Arbeitstitel)*.
- Storper, M. (1995). The Resurgence of Regional Economics, Ten Years Later: The Region as a Nexus of Untraded Interdependencies. *European Urban and Regional Studies*, 2(3), 191–221.
- Sun, L.-Y., Aryee, S. & Law, K. S. (2007). High-Performance Human Resource Practices, Citizenship Behaviour, and Organizational Performance: A Relational Perspective. *Academy of Management Journal*, 50(3), 558–577.
- Sundaram, A. K. & Inkpen, A. C. (2004a). Stakeholder Theory and "The Corporate Objective Revisited": A Reply. *Organization Science*, 15(3), 370–371.
- Sundaram, A. K. & Inkpen, A. C. (2004b). The Corporate Objective Revisited. *Organization Science*, 15(3), 350–363.
- Szulanski, G. & Jensen, R. J. (2006). Presumptive Adaption and the Effectiveness of Knowledge Transfer. *Strategic Management Journal*, 27(10), 937–957.
- Takeuchi, R., Shay, J. P. & Li, J. (2008). When Does Decision Autonomy Increase Expatriate Managers' Adjustment? An Empirical Test. *Academy of Management Journal*, 51(1), 45–60.
- Teece, D. J. (2000). Strategies for Managing Knowledge Assets: The Role of Firm Structure and Industrial Context. *Long Range Planning*, 33, 35–54.
- Teece, D. J., Pisano, G. & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7), 509–533.

- Terlaak, A. & King, A. A. (2007). Follow the Small? Information-Revealing Adoption Bandwagons when Observers Expect Larger Firms to Benefit More from Adoption. *Strategic Management Journal*, 28(12), 1167–1185.
- Thommen, J.-P. & Achleitner, A.-K. (2006). *Allgemeine Betriebswirtschaftslehre: Umfassende Einführung aus managementorientierter Sicht* (5., überarb. und erw. Aufl., [Hauptbd.]). Wiesbaden: Gabler.
- Thornhill, S. & White, R. E. (2007). Strategic Purity: A Multi-Industry-Evaluation of Pure Vs. Hybrid Business Strategies. *Strategic Management Journal*, 28(5), 553–561.
- Tirole, J. (1999). Incomplete Contracts: Where Do We Stand? *Econometrica*, 67(4), 741–781.
- Tong, T. W., Reuer, J. J. & Peng, M. W. (2008). International Joint Ventures and the Value of Growth Options. *Academy of Management Journal*, 51(5), 1014–1029.
- Tversky, A. & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232.
- U. S. Census Bureau (2007). *Number of Firms, Number of Establishments, Employment, and Annual Payroll by Employment Size of the Enterprise for the United States and States, Totals - 2005*. Retrieved July 30, 2008, from <http://www2.census.gov/csd/susb/2005/usst05.xls>.
- Uhlenbruck, K., Hitt, M. A. & Semadeni, M. (2006). Market Value Effects of Acquisitions Involving Internet Firms: A Resource-based Analysis. *Strategic Management Journal*, 27(10), 899–913.
- Uotila, J., Maula, M., Keil, T. & Zahra, S. A. (2009). Exploration, Exploitation, and Financial Performance: Analysis of S&P 500 Corporations. *Strategic Management Journal*, 30(2), 221–231.
- Venkatraman, N. & Ramanujam, V. (1986). Measurement of Business Performance in Strategic Research: A Comparison of Approaches. *Academy of Management Review*, 11(4), 801–814.
- Vroom, G. & Gimeno, J. (2007). Ownership Form, Managerial Incentives, and the Intensity of Rivalry. *Academy of Management Journal*, 50(4), 901–922.
- Wadhwa, A. & Kotha, S. (2006). Knowledge Creation Through External Venturing: Evidence from the Telecommunications Equipment Manufacturing Industry. *Academy of Management Journal*, 49(4), 819–835.
- Wall, T. D., Michie, J., Patterson, M., Wood, S. J., Sheehan, M. & Clegg, C. W. (2004). On The Validity of Subjective Measures of Company Performance. *Personnel Psychology*, 57(1), 95–118.
- Wan, W. P. & Yiu, D. W. (2009). From Crisis to Opportunity: Environmental Jolt, Corporate Acquisitions, and Firm Performance. *Strategic Management Journal*, 30(7), 791–801.
- Wang, L. & Zajac, E. J. (2007). Alliance or Acquisition? A Dyadic Perspective on Interfirm Resource Combinations. *Strategic Management Journal*, 28(13), 1291–1317.
- Wasserman, N. (2006). Stewards, Agents, and the Founder Discount: Executive Compensation in New Ventures. *Academy of Management Journal*, 49(5), 960–976.
- Weigelt, C. (2009). The Impact of Outsourcing New Technologies on Integrative Capabilities and Performance. *Strategic Management Journal*, 30(6), 595–616.
- Weigelt, C. & Sarkar, M. B. (2009). Learning from Supply-Side Agents: The Impact of Technology Solution Providers' Experiential Diversity on Clients' Innovation Adoption. *Academy of Management Journal*, 52(1), 37–60.
- Weik, E. & Lang, R. (Eds.) (2005). *Moderne Organisationstheorien 1: Handlungsorientierte Ansätze* (2., überarb. Aufl., Vol. 1). Wiesbaden: Gabler.

- Westphal, J. D. & Stern, I. (2007). Flattery Will Get You Everywhere (Especially If You Are a Male Caucasian): How Integration, Boardroom Behaviour, and Demographic Minority Status Affect Additional Board Appointments at U.S. Companies. *Academy of Management Journal*, 50(2), 267–288.
- Wiersema, M. F. & Bowen, H. P. (2009a). Corporate Diversification: The Impact of Foreign Competition, Industry Globalization, and Product Diversification. *Strategic Management Journal*, 28(2), 115–132.
- Wiersema, M. F. & Bowen, H. P. (2009b). The Use of Limited Dependent Variable Techniques in Strategy Research: Issues and Methods. *Strategic Management Journal*, 30(6), 679–692.
- Wiggins, R. R. & Ruefli, T. W. (2002). Sustained Competitive Advantage: Temporal Dynamics and the Incidence and Persistence of Superior Economic Performance. *Organization Science*, 13(1), 82–105.
- Williams, J. R. (1994). Strategy and the Search for Rents: The Evolution of Diversity among Firms. In R. P. Rumelt; D. Schendel & D. J. Teece (Eds.), *Fundamental Issues in Strategy: A Research Agenda* (pp. 229–246). Boston: Harvard Business School Press.
- Williams, L. J., Gavin, M. B. & Hartman, N. S. (2004). Structural Equation Modeling Methods in Strategy Research: Applications and Issues. In D. J. Ketchen & D. D. Bergh (Eds.), *Research Methodology in Strategy and Management* (Vol. 1, pp. 303–346). Amsterdam u.a.: Elsevier.
- Williamson, O. E. (1999). Strategy Research: Governance and Competence Perspectives. *Strategic Management Journal*, 20(12), 1087–1108.
- Wolf, J. (2008). *Organisation, Management, Unternehmensführung: Theorien, Praxisbeispiele und Kritik* (3., vollständig überarbeitete und erweiterte Auflage). Wiesbaden: Gabler Verlag / GWV Fachverlage GmbH Wiesbaden.
- Wortman, P. M. (1994). Judging Research Quality. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 97–110). New York: Russell Sage Foundation.
- Wright, P., Kroll, M., Krug, J. A. & Pettus, M. (2007). Influences of Top Management Team Incentives on Firm Risk Taking. *Strategic Management Journal*, 28(1), 81–89.
- Zahra, S. A., Ireland, R. Duane & Hitt, M. A. (2000). International Expansion by New Venture Firms: International Diversity, Mode of Market Entry, Technological Learning, and Performance. *Academy of Management Journal*, 43(5), 925–950.
- Zahra, S. A. & Nielsen, A. P. (2002). Sources of Capabilities, Integration and Technology Commercialization. *Strategic Management Journal*, 23(5), 377–398.
- Zapf, D., Dormann, C. & Frese, M. (1996). Longitudinal Studies in Organizational Stress Research: A Review of the Literature With References to Methodological Issues. *Journal of Occupational Health Psychology*, 1(2), 145–169.
- Zhang, X., Bartol, K. M., Smith, K. G., Pfarrer, M. D. & Khanin, D. M. (2008). CEOs on the Edge: Earnings Manipulation and Stock-Based Incentive Misalignment. *Academy of Management Journal*, 51(2), 241–258.
- Zhou, K. Zheng, Li, J. Juan, Zhou, N. & Su, C. (2008). Market Orientation, Job Satisfaction, Product Quality, and Firm Performance: Evidence from China. *Strategic Management Journal*, 29(9), 985–1000.
- Zimmerman, J. L. (2001). Conjectures regarding empirical managerial accounting research. *Journal of Accounting and Economics*, 32, 411–427.
- Zingales, L. (2000). In Search of New Foundations. *Journal of Finance*, 55(4), 1623–1653.
- Zott, C. & Amit, R. (2008). The Fit Between Product Market Strategy and Business Modell: Implications for Firm Performance. *Strategic Management Journal*, 29(1), 1–26.

Internetquellen

<http://bea.gov> – Das Bureau of Economic Analysis, U. S. Department of Commerce, letzter Zugriff am 24.11.2009.

www.destatis.de – Das Statistische Bundesamt, letzter Zugriff am 30.7.2009.

www.deutsche-boerse.com – Der Internet-Auftritt der Deutsche Börse Group, letzter Zugriff am 30.7.2009.

www.dialog-erfolgsfaktorenforschung – Die Diskussion von Kieser und Nicolai sowie Widerreden zu Sinn und Unsinn der Erfolgsfaktorenforschung, letzter Zugriff am 17.07.2009.

www.forschungsmonitoring.org – Das Forschungsmonitoring des Vereins für Socialpolitik, letzter Zugriff am 30.07.2009.

<http://journals.aomonline.org/amj/home.html> - Das Academy of Management Journal, letzter Zugriff am 23.3.2009.

www.lexikon.meyers.de – Die Online-Ausgabe von Meyers Universallexikon, letzter Zugriff am 28.10.2008.

<http://www.nyse.com/> - Der Internet-Auftritt der NYSE Euronext, letzter Zugriff am 3.08.2009.

<http://strategicmanagement.net/pub/smj.php> – Das Strategic Management Journal, letzter Zugriff am 10.12.2008.

www.zeit.de – Der Online-Auftritt der Wochenzeitung „Die Zeit“, letzter Zugriff am 28.10.2008.

Literaturverzeichnis II: Verzeichnis der in die qualitative Review einbezogenen Artikel

- Acquaah, M. (2007). Managerial Social Capital, Strategic Orientation, and Organizational Performance in an Emerging Economy. *Strategic Management Journal*, 28(12), 1235–1255.
- Aggarwal, V. A. & Hsu, D. H. (2009). Modes of Cooperative Commercialization by Start-Ups. *Strategic Management Journal*, 30(8), 835–864.
- Andersen, T. J., Denrell, J. & Bettis, R. A. (2007). Strategic Responsiveness and Bowman's Risk-Return Paradox. *Strategic Management Journal*, 28(4), 407–429.
- Ang, S. Hwee (2008). Competitive Intensity and Collaboration: Impact on Firm Growth Across Technological Environments. *Strategic Management Journal*, 29(10), 1057–1075.
- Arthaud-Day, M. L., Certo, S. Trevis, Dalton, C. M. & Dalton, D. R. (2006). A Changing of the Guard: Executive and Director Turnover Following Corporate Financial Restatements. *Academy of Management Journal*, 49(6), 1119–1136.
- Barden, J. Q. & Mitchell, W. (2007). Disentangling the Influences of Leaders' Relational Embeddedness on Interorganizational Exchange. *Academy of Management Journal*, 50(6), 1440–1461.
- Barkema, H. G. & Schijven, M. (2008). Toward Unlocking the Full Potential of Acquisitions: The Role of Organizational Restructuring. *Academy of Management Journal*, 51(4), 696–722.
- Barnett, M. L. & King, A. A. (2008). Good Fences Make Good Neighbors: A Longitudinal Analysis of an Industry Self-Regulatory Institution. *Academy of Management Journal*, 51(6), 1150–1170.
- Barnett, M. L. & Salomon, R. M. (2006). Beyond Dichotomy: The Curvilinear Relationship Between Social Responsibility and Financial Performance. *Strategic Management Journal*, 27(11), 1101–1122.
- Barrik, M. R., Bradley, B. H., Kristof-Brown, A. L. & Colbert, A. E. (2007). The Moderating Role of Top Management Team Interdependence: Implications for Real Teams and Working Groups. *Academy of Management Journal*, 50(3), 544–557.
- Basdeo, D. K., Smith, K. G., Grimm, C. M., Rindova, V. P. & Derfus, P. J. (2006). The Impact of Market Actions on Firm Reputation. *Strategic Management Journal*, 27(12), 1205–1219.
- Beckman, C. M. (2006). The Influence of Founding Team Company Affiliations on Firm Behaviour. *Academy of Management Journal*, 49(4), 741–758.
- Bercovitz, J. & Mitchell, W. (2007). When is More Better? The Impact of Business Scale and Scope on Long-Term Business Survival, While Controlling for Profitability. *Strategic Management Journal*, 28(1), 61–79.
- Bergh, D. D., Johnson, R. A. & Dewitt, R.-L. (2008). Restructuring Through Spin-Off or Sell-Off: Transforming Information Asymmetries Into Financial Gain. *Strategic Management Journal*, 29(2), 133–148.
- Bergh, D. D. & Ngah-Kiing Lim, E. (2008). Learning how to Restructure: Absorptive Capacity and Improvisational Views of Restructuring Actions and Performance. *Strategic Management Journal*, 29(6), 593–616.
- Berrone, P. & Gomez-Mejia, L. R. (2009). Environmental Performance and Executive Compensation: An Integrated Agency-Institutional Perspective. *Academy of Management Journal*, 52(1), 103–126.
- Berry, H. (2006). Shareholder Valuation of Foreign Investment and Expansion. *Strategic Management Journal*, 27(12), 1123–1140.

- Bonardi, J.-P., Holburn, G. L. F. & Vanden Bergh, R. G. (2006). Nonmarket Strategy Performance: Evidence from U.S. Electric Utilities. *Academy of Management Journal*, 49(6), 1209–1228.
- Bouquet, C. & Birkinshaw, J. (2008). Weight Versus Voice: How Foreign Subsidiaries Gain Attention from Corporate Headquarters. *Academy of Management Journal*, 51(3), 577–601.
- Boyd, J. L. & Bresser, R. K. F. (2008). Performance implications of delayed competitive responses: evidence from the US retail industry. *Strategic Management Journal*, 29(10), 1077–1096.
- Brammer, S. & Millington, A. (2008). Does it Pay to be Different? An Analysis of the Relationship Between Corporate Social and Financial Performance. *Strategic Management Journal*, 29(12), 1325–1343.
- Buo, J. Carlos & Satorra, A. (2007). The Persistence of Abnormal Returns at Industry and Firm Levels: Evidence from Spain. *Strategic Management Journal*, 28(7), 707–722.
- Burgelman, R. A. & Grove, A. S. (2007). Let Chaos Reign, Then Rein in Chaos - Repeatedly: Managing Strategic Dynamics for Corporate Longevity. *Strategic Management Journal*, 28(10), 965–979.
- Canella Jr., A. A., Park, J.-H. & Lee, H.-U. (2008). Top Management Team Functional Background Diversity and Firm Performance: Examining the Role of Team Member Colocation and Environmental Uncertainty. *Academy of Management Journal*, 51(4), 768–784.
- Capaldo, A. (2007). Network Structure and Innovation: The Leveraging of a Dual Network as a Distinctive Relational Capability. *Strategic Management Journal*, 28(6), 585–608.
- Capron, L. & Shen, J.-C. (2007). Acquisitions of Private Vs. Public Firms: Private Information, Target Selection, and Acquirer Returns. *Strategic Management Journal*, 28(9), 891–911.
- Ceccagnoli, M. (2009). Appropriability, Preemption, and Firm Performance. *Strategic Management Journal*, 30(1), 81–98.
- Chakrabarti, A., Singh, K. & Mahmood, I. (2007). Diversification and Performance: Evidence from East Asian Firms. *Strategic Management Journal*, 28(2), 101–120.
- Chan, C. M., Isobe, T. & Makino, S. (2008). Which Country Matters? Institutional Development and Foreign Affiliate Performance. *Strategic Management Journal*, 29(11), 1179–1205.
- Chang, S. Jin & Xu, D. (2008). Spillovers and Competition Among Foreign and Local Firms in China. *Strategic Management Journal*, 29(11), 495–518.
- Chatterji, A. K. (2009). Spawned with a Silver Spoon?: Entrepreneurial Performance and Innovation in the Medical Device Industry. *Strategic Management Journal*, 30(2), 185–206.
- Chen, M.-J., Su, K.-H. & Tsai, W. (2007). Competitive Tension: The Awareness-Motivation-Capability Perspective. *Academy of Management Journal*, 50(1), 101–118.
- Chen, W.-R. & Miller, K. D. (2007). Situational and Institutional Determinants of Firm R&D Search Intensity. *Strategic Management Journal*, 28(4), 369–381.
- Cho, T. S. & Shen, W. (2007). Changes in Executive Compensation Following an Environmental Shift: The Role of Top Management Team Turnover. *Strategic Management Journal*, 28(7), 747–754.
- Choi, J. & Wang, H. (2009). Stakeholder Relations and the Persistence of Corporate Financial Performance. *Strategic Management Journal*, 30(8), 895–907.
- Colbert, A. E., Kristof-Brown, A. L., Bradley, B. H. & Barrik, M. R. (2008). CEO Transformational Leadership: The Role of Goal Importance Congruence in Top Management Teams. *Academy of Management Journal*, 51(1), 81–96.
- Collis, D., Young, D. & Goold, M. (2007). The Size, Structure, and Performance of Corporate Headquarters. *Strategic Management Journal*, 28(4), 383–405.

- Cording, M., Christmann, P. & King, D. R. (2008). Reducing Causal Ambiguity in Acquisition Integration: Intermediate Goals as Mediators of Integration Decisions and Acquisition Performance. *Academy of Management Journal*, 51(4), 744–767.
- Crook, T. Russell, Ketchen, D. J., Combs, J. G. & Todd, S. Y. (2008). Strategic Resources and Performance: A Meta-Analysis. *Strategic Management Journal*, 29(11), 1141–1154.
- David, P., Bloom, M. & Hillman, A. J. (2007). Investor Activism, Managerial Responsiveness, and Corporate Social Performance. *Strategic Management Journal*, 28(1), 91–100.
- David, P., O'Brien, J. P. & Yoshikawa, T. (2008). The Implications of Dept Heterogeneity for R&D Investment and Firm Performance. *Academy of Management Journal*, 51(1), 165–181.
- Derfus, P. J., Maggitti, P. G., Grimm, C. M. & Smith, K. G. (2008). The Red Queen Effect: Competitive Actions and Firm Performance. *Academy of Management Journal*, 51(1), 61–80.
- Devers, C. E., Wiseman, R. M. & Holmes Jr., R. Michael (2007). The Effects of Endowment and Loss Aversion in Managerial Stock Option Valuation. *Academy of Management Journal*, 50(1), 191–208.
- Dowell, G. (2006). Product Line Strategies of New Entrants in an Established Industry: Evidence from the U.S. Bicycle Industry. *Strategic Management Journal*, 27(10), 959–979.
- Dowell, G. & Swaminathan, A. (2006). Entry Timing, Exploration and Firm Survival in the Early U.S. Bicycle Industry. *Strategic Management Journal*, 27(12), 1159–1182.
- Durand, R., Rao, H. & Monin, P. (2007). Code and Conduct in French Cuisine: Impact of Code Changes on External Evaluations. *Strategic Management Journal*, 28(5), 455–472.
- Eisenmann, T. R. (2006). Internet Companies' Growth Strategies: Determinants of Investment Intensity and Long-Term Performance. *Strategic Management Journal*, 27(12), 1183–1204.
- Elbanna, S. & Child, J. (2007). Influences on Strategic Decision Effectiveness: Development and Test of an Integrative Model. *Strategic Management Journal*, 28(4), 431–453.
- Ellis, K. M., Reus, T. H. & Lamont, B. T. (2009). The Effects of Procedural and Informational Justice in the Integration of Related Acquisitions. *Strategic Management Journal*, 30(2), 137–161.
- Ethiraj, S. & Zhu, D. H. (2008). Performance Effects of Imitative Entry. *Strategic Management Journal*, 29(8), 797–817.
- Fang, Y., Wade, M. & Delios, A. (2007). International Diversification, Subsidiary Performance, and the Mobility of Knowledge Resources. *Strategic Management Journal*, 28(10), 1053–1064.
- Fiss, P. C. (2006). Social Influence Effects and Managerial Compensation: Evidence from Germany. *Strategic Management Journal*, 27(11), 1013–1031.
- Fiss, P. C. & Zajac, E. J. (2006). The Symbolic Management of Strategic Change: Sensegiving via Framing and Decoupling. *Academy of Management Journal*, 49(6), 1173–1193.
- Fitza, M., Matusik, S. F. & Mosakowski, E. (2009). Do VCs Matter? The Importance of Owners on Performance Variance in Start-Up Firms. *Strategic Management Journal*, 30(4), 387–404.
- Godfrey, P. C., Merrill, C. B. & Hansen, J. M. (2009). The Relationship between Corporate Social Responsibility and Shareholder Value: An Empirical Test of the Risk Management Hypothesis. *Strategic Management Journal*, 30(4), 425–445.
- Goerzen, A. (2007). Alliance Networks and Firm Performance: The Impact of Repeated Partnerships. *Strategic Management Journal*, 28(5), 487–509.

- Gong, Y., Shenkar, O., Luo, Y. & Nyaw, M.-K. (2007). Do Multiple Parents Help or Hinder International Joint Venture Performance? The Mediating Role of Contract Completeness and Partner Cooperation. *Strategic Management Journal*, 28(10), 1021–1034.
- Goranova, M., Alessandri, T. M., Brandes, P. & Dharwadkar, R. (2007). Managerial Ownership and Corporate Diversification: A Longitudinal View. *Strategic Management Journal*, 28(3), 211–225.
- Greve, H. R. (2008a). A Behavioral Theory of Firm Growth: Sequential Attention to Size and Performance Goals. *Academy of Management Journal*, 51(3), 476–494.
- Greve, H. R. (2008b). Multimarket Contact and Sales Growth: Evidence from Insurance. *Strategic Management Journal*, 29(3), 229–249.
- Haas, M. R. & Hansen, M. T. (2007). Different Knowledge, Different Benefits: Toward a Productivity Perspective on Knowledge Sharing in Organizations. *Strategic Management Journal*, 28(11), 1133–1153.
- Heeley, M. B. & Jacobson, R. (2008). The Recency of Technological Inputs and Financial Performance. *Strategic Management Journal*, 29(7), 723–744.
- Henkel, J. (2009). The Risk-Return Paradox for Strategic Management: Disentangling True and Spurious Effects. *Strategic Management Journal*, 30(3), 287–303.
- Heugens, P. P. M. A. R. & Lander, M. W. (2009). Structure! Agency! (And Other Quarrels): A Meta-Analysis of Institutional Theories of Organizations. *Academy of Management Journal*, 52(1), 61–85.
- Hillman, A. J., Shropshire, C. & Canella Jr., A. A. (2007). Organizational Predictors of Woman on Corporate Boards. *Academy of Management Journal*, 50(4), 941–952.
- Hitt, M. A., Bierman, L., Uhlenbruck, K. & Shimizu, K. (2006). The Importance of Resources in the Internationalization of Professional Service Firms: the Good, the Bad, and the Ugly. *Academy of Management Journal*, 49(6), 1137–1157.
- Hmieleski, K. M. & Baron, R. A. (2009). Entrepreneurs' Optimism and New Venture Performance: A Social Cognitive Perspective. *Academy of Management Journal*, 52(3), 473–488.
- Holcomb, T. R., Holmes Jr., R. Michael & Connelly, B. L. (2009). Making the Most of What You Have: Managerial Ability as a Source of Resource Value Creation. *Strategic Management Journal*, 30(5), 457–485.
- Huckman, R. S. & Zinner, D. E. (2008). Does Focus Improve Operational Performance?: Lessons from the Management of Clinical Trials. *Strategic Management Journal*, 29(2), 173–193.
- Hull, C. Eirikur & Rothenberg, S. (2008). Firm Performance: The Interactions of Corporate Social Performance with Innovation and Industry Differentiation. *Strategic Management Journal*, 29(7), 781–789.
- Hult, G. Thomas M., Ketchen, D. J. & Arrfelt, M. (2007). Strategic Supply Chain Management: Improving Performance Through a Culture of Competitiveness and Knowledge Development. *Strategic Management Journal*, 28(10), 1035–1052.
- Iyer, D. & Miller, K. D. (2008). Performance Feedback, Slack, and the Timing of Acquisitions. *Academy of Management Journal*, 51(4), 808–822.
- Jensen, M. & Roy, A. (2008). Staging Exchange Partner Choices: When do Status and Reputation Matter? *Academy of Management Journal*, 51(3), 495–516.
- Julian, S. D. & Ofori-Dankwa, J. C. (2008). Toward an Integrative Cartography of Two Strategic Issues Diagnosis Frameworks. *Strategic Management Journal*, 29(1), 93–114.
- Kabanoff, B. & Brown, S. (2008). Knowledge Structures of Prospectors, Analyzers, and Defenders: Content, Structure, Stability, and Performance. *Strategic Management Journal*, 29(2), 149–171.

- Kacperczyk, A. (2009). With Greater Power Comes Greater Responsibility? Takeover Protection and Corporate Attention to Stakeholders. *Strategic Management Journal*, 30(3), 261–285.
- Kale, P. & Singh, H. (2007). Building Firm Capabilities Through Learning: The Role of the Alliance Learning Process in Alliance Capability and Firm-Level Alliance Success. *Strategic Management Journal*, 28(10), 981–1000.
- Kalyta, P. (2009). Compensation Transparency and Managerial Opportunism: A Study of Supplemental Retirement Plans. *Strategic Management Journal*, 30(4), 405–423.
- Kang, E. (2008). Director Interlocks and Spillover Effects of Reputational Penalties from Financial Reporting Fraud. *Academy of Management Journal*, 51(3), 537–555.
- Kapoor, R. & Lim, K. (2007). The Impact of Acquisitions on the Productivity of Inventors at Semiconductor Firms: A Synthesis of Knowledge-Based and Incentive-Based Perspectives. *Academy of Management Journal*, 50(5), 1133–1155.
- Karaevli, A. (2007). Performance Consequences of New CEO 'Outsiderness': Moderating Effects of Pre- and Post-Succession Contexts. *Strategic Management Journal*, 28(7), 681–706.
- Karnani, A. (2007). Doing Well by Doing Good - Case Study: "Fair & Lovely" Whitening Cream. *Strategic Management Journal*, 28(13), 1351–1357.
- Katsikeas, C. S., Samiee, S. & Theodosiou, M. (2006). Strategy Fit and Performance Consequences of International Marketing Standardization. *Strategic Management Journal*, 27(9), 867–890.
- Keil, T., Maula, M., Schildt, H. & Zhara, S. A. (2008). The Effect of Governance Modes and Relatedness of External Business Development Activities on Innovative Performance. *Strategic Management Journal*, 29(8), 895–907.
- Kim, J.-Y. (Jay) & Finkelstein, S. (2009). The Effects of Strategic and Market Complementarity on Acquisition Performance: Evidence from the U.S. Commercial Banking Industry, 1989–2001. *Strategic Management Journal*, 30(6), 617–646.
- Kim, J.-Y. (Jay) & Miner, A. S. (2007). Vicarious Learning From the Failures and Near-Failures of Others: Evidence from the U.S. Commercial Banking Industry. *Academy of Management Journal*, 50(3), 687–714.
- Koka, B. R. & Prescott, J. E. (2008). Designing alliance networks: the influence of network position, environmental change, and strategy on firm performance. *Strategic Management Journal*, 29(6), 639–661.
- Kor, Y. Y. (2006). Direct and Interaction Effects of Top Management Team and Board Compositions on R&D Investment Strategy. *Strategic Management Journal*, 27(11), 1081–1099.
- Krishnan, R., Martin, X. & Noorderhaven, N. G. (2006). When Does Trust Matter to Alliance Performance? *Academy of Management Journal*, 49(5), 894–917.
- Kroll, M., Walters, B. A. & Le, S. A. (2007). The Impact of Board Composition and Top Management Team Ownership Structure on Post-IPO Performance in Young Entrepreneurial Firms. *Academy of Management Journal*, 50(5), 1198–1216.
- Kroll, M., Walters, B. A. & Wright, P. (2008). Board Vigilance, Director Experience, and Corporate Outcomes. *Strategic Management Journal*, 29(4), 363–382.
- Kronborg, D. & Thomsen, S. (2009). Foreign Ownership and Long-Term Survival. *Strategic Management Journal*, 30(2), 207–219.
- Kumar, M. V. Shyam (2009). The Relationship Between Product and International Diversification: The Effects of Short-Run Constraints and Endogeneity. *Strategic Management Journal*, 30(1), 99–1116.

- Laamanen, T. (2007). On the Role of Acquisition Premium in Acquisition Research. *Strategic Management Journal*, 28(13), 1359–1369.
- Laamanen, T. & Keil, T. (2008). Performance of Serial Acquirers: Toward an Acquisition Program Perspective. *Strategic Management Journal*, 29(6), 663–672.
- Lado, A. A., Dant, R. R. & Tekleab, A. G. (2008). Trust-Opportunism Paradox, Relationalism, and Performance in Interfirm Relationships: Evidence from the Retail Industry. *Strategic Management Journal*, 29(4), 401–423.
- Lamberg, J.-A., Tikkanen, H., Nokelainen, T. & Suur-Inkeroinen, H. (2009). Competitive Dynamics, Strategic Consistency, and Organizational Survival. *Strategic Management Journal*, 30(1), 45–60.
- Lange, D., Boivie, S. & Henderson, A. D. (2009). The Parenting Paradox: How Multi-Business Diversifiers Endorse Disruptive Technologies While Their Corporate Children Struggle. *Academy of Management Journal*, 52(1), 179–198.
- Larraza-Kintana, M., Wiseman, R. M., Gomez-Mejia, L. R. & Welbourne, T. M. (2007). Disentangling Compensation and Employment Risks Using the Behavioral Agency Model. *Strategic Management Journal*, 28(10), 1001–1019.
- Lavie, D. (2007). Alliance Portfolios and Firm Performance: A Study of Value Creation and Appropriation in the U.S. Software Industry. *Strategic Management Journal*, 28(12), 1187–1212.
- Lavie, D., Lechner, C. & Singh, H. (2007). The Performance Implications of Timing of Entry and Involvement in Multipartner Alliances. *Academy of Management Journal*, 50(3), 578–604.
- Lavie, D. & Rosenkopf, L. (2006). Balancing Exploration and Exploitation in Alliance Formation. *Academy of Management Journal*, 49(4), 797–818.
- Lazzarini, S. G. (2007). The Impact of Membership in Competing Alliance Constellations: Evidence on the Operational Performance of Global Airlines. *Strategic Management Journal*, 28(4), 345–367.
- Leask, G. & Parker, D. (2007). Strategic Groups, Competitive Groups and Performance Within the U.K. Pharmaceutical Industry: Improving Our Understanding of the Competitive Process. *Strategic Management Journal*, 28(7), 723–745.
- Lee, P. M. & James, E. Hayes (2007). She'-E-Os: Gender Effects and Investors Reactions to the Announcements of Top Executive Appointments. *Strategic Management Journal*, 28(3), 227–241.
- Lee, S.-H. & Makhija, M. (2009). Flexibility in Internationalization: Is it Valuable During an Economic Crisis? *Strategic Management Journal*, 30(5), 537–555.
- Leiblein, M. J. & Madsen, T. L. (2009). Unbundling Competitive Heterogeneity: Incentive Structures and Capability Influences on Technological Innovation. *Strategic Management Journal*, 30(7), 711–735.
- Leiponen, A. (2008). Control of Intellectual Assets in Client Relationships: Implications for Innovation. *Strategic Management Journal*, 29(13), 1371–1394.
- Levitas, E. & McFayden, M. Ann (2009). Managing Liquidity in Research-Intensive Firms: Signaling and Cash-Flow Effects of Patents and Alliance Activities. *Strategic Management Journal*, 30(6), 659–678.
- Li, J., Sun, J. J. M. & Liu, S. X. Y. (2008). Strategic Human Resource Management, Institutionalization and Employment Modes: An Empirical Study in China. *Strategic Management Journal*, 29(3), 337–342.
- Li, J., Zhou, C. & Zajac, E. J. (2009). Control, Collaboration and Productivity in International Joint Ventures: Theory and Evidence. *Strategic Management Journal*, 30(8), 865–884.

- Li, J. Juan, Poppo, L. & Zhou, K. Zheng (2008). Do Managerial Ties in China Always Produce Value? Competition, Uncertainty, and Domestic vs. Foreign Firms. *Strategic Management Journal*, 29(4), 383–400.
- Ling, Y., Simsek, Z., Lubatkin, M. & Veiga, J. F. (2008). Transformational Leadership's Role in Promoting Corporate Entrepreneurship: Examining the CEO-TMT Interface. *Academy of Management Journal*, 51(3), 557–576.
- Love, E. Geoffrey & Kraatz, M. (2009). Character, Conformity, or the Bottom Line? How and Why Downsizing Affected Corporate Reputation. *Academy of Management Journal*, 52(2), 314–335.
- Lu, J. W. & Ma, X. (2008). The Contingent Value of Local Partners' Business Group Affiliations. *Academy of Management Journal*, 51(2), 295–314.
- Lunnan, R. & Haugland, S. A. (2008). Predicting and Measuring Alliance Performance: A Multi-Dimensional Analysis. *Strategic Management Journal*, 29(5), 545–556.
- Luo, Y. (2007a). Are Joint Venture Partners More Opportunistic in a More Volatile Environment? *Strategic Management Journal*, 28(1), 39–60.
- Luo, Y. (2007b). The Independent and Interactive Roles of Procedural, Distributive, and Interactional Justice in Strategic Alliances. *Academy of Management Journal*, 50(3), 644–664.
- Luo, Y. (2008a). Procedural Fairness and Interfirm Cooperation in Strategic Alliances. *Strategic Management Journal*, 29(1), 27–46.
- Luo, Y. (2008b). Structuring Interorganizational Cooperation: The Role of Economic Integration in Strategic Alliances. *Strategic Management Journal*, 29(6), 617–637.
- Macher, J. T. & Boerner, C. S. (2006). Experience and Scale and Scope Economies: Trade-Offs and Performance in Development. *Strategic Management Journal*, 27(9), 845–865.
- Mackey, A. (2008). The Effect of CEOs on Firm Performance. *Strategic Management Journal*, 29(12), 1357–1367.
- Makri, M., Lane, P. J. & Gomez-Mejia, L. R. (2006). CEO Incentives, Innovation, and Performance in Technology-Intensive Firms: A Reconciliation of Outcome and Behavior-Based Incentive Schemes. *Strategic Management Journal*, 27(11), 1057–1080.
- Marcel, J. J. (2009). Why Top Management Team Characteristics Matter When Employing a Chief Operating Officer: A Strategic Contingency Perspective. *Strategic Management Journal*, 30(6), 647–658.
- McDonald, M. L., Khanna, P. & Westphal, J. D. (2008). Getting Them to Think Outside the Circle: Corporate Governance, CEOs' External Advice Networks, and Firm Performance. *Academy of Management Journal*, 51(3), 453–475.
- McDonald, M. L., Westphal, J. D. & Graebner, M. E. (2008). What Do They Know? The Effects of Outside Director Acquisition Experience on Firm Acquisition Performance. *Strategic Management Journal*, 29(11), 1155–1177.
- McNamara, G. M., Haleblan, J. (John) & Dykes, B. Johnson (2008). The Performance Implications of Participating in an Acquisition Wave: Early Mover Advantages, Bandwagon Effects, and the Moderating Influence of Industry Characteristics and Acquirer Tactics. *Academy of Management Journal*, 51(1), 113–130.
- Mesquita, L. F., Anand, J. & Brush, T. H. (2008). Comparing the Resource-based and Relational View: Knowledge Transfer and Spillover in Vertical Alliances. *Strategic Management Journal*, 29(9), 913–941.
- Mesquita, L. F. & Brush, T. H. (2008). Untangling Safeguard and Production Coordination Effects in Long-Term Buyer-Supplier Relationships. *Academy of Management Journal*, 51(4), 785–807.

- Mesquita, L. F. & Lazzarini, S. G. (2008). Horizontal and Vertical Relationships in Developing Economies: Implications for SMEs' Access to Global Markets. *Academy of Management Journal*, 51(2), 359–380.
- Moliterno, T. P. & Wiersema, M. F. (2007). Firm Performance, Rent Appropriation, and the Strategic Resource Divestment Capability. *Strategic Management Journal*, 28(11), 1065–1087.
- Morrow Jr., J. L., Sirmon, D. G., Hitt, M. A. & Holcomb, T. R. (2007). Creating Value in the Face of Declining Performance: Firm Strategies and Organizational Recovery. *Strategic Management Journal*, 28(3), 271–283.
- Nadkarni, S. & Barr, P. S. (2008). Environmental Context, Managerial Cognition, and Strategic Action: An Integrated View. *Strategic Management Journal*, 29(13), 1395–1427.
- Nadkarni, S. & Narayanan, V. K. (2007). Strategic Schemas, Strategic Flexibility, and Firm Performance: The Moderating Role of Industry Clockspeed. *Strategic Management Journal*, 28(3), 243–270.
- Newbert, S. L. (2007). Empirical Research on the Resource-Based View of the Firm: An Assessment and Suggestions for Future Research. *Strategic Management Journal*, 28(2), 121–146.
- Newbert, S. L. (2008). Value, Rareness, Competitive Advantage, and Performance: A Conceptual-Level Empirical Investigation of the Resource-based View of the Firm. *Strategic Management Journal*, 29(7), 745–768.
- Ngah-Kiing Lim, E., Das, S. S. & Das, A. (2009). Diversification Strategy, Capital Structure, and the Asian Financial Crisis (1997-1998): Evidence from Singapore Firms. *Strategic Management Journal*, 30(6), 577–594.
- Nordenflycht, A. von (2007). Is Public Ownership Bad for Professional Service Firms? Ad Agency Ownership, Performance, and Creativity. *Academy of Management Journal*, 50(2), 429–445.
- Oriani, R. & Sobrero, M. (2008). Uncertainty and the Market Valuation of R&D Within a Real Options Logic. *Strategic Management Journal*, 29(4), 343–361.
- Ozcan, P. & Eisenhardt, K. M. (2009). Origin of Alliance Portfolios: Entrepreneurs, Network Strategies, and Firm Performance. *Academy of Management Journal*, 52(2), 246–279.
- Pacheco-de-Almeida, G., Henderson, J. E. & Cool, K. O. (2008). Resolving the Commitment versus Flexibility Trade-Off: The Role of Resource Accumulation Lags. *Academy of Management Journal*, 51(3), 517–536.
- Peteraf, M. A. & Reed, R. (2007). Managerial Discretion and Internal Alignment under Regulatory Constraints and Change. *Strategic Management Journal*, 28(11), 1089–1112.
- Pierce, L. (2009). Big Losses in Ecosystem Niches: How Core Firm Decisions Drive Complementary Product Shakeouts. *Strategic Management Journal*, 30(3), 323–347.
- Prince, J. T. & Simon, D. H. (2009). Multimarket Contact and Service Quality: Evidence from On-Time Performance in the U.S. Airline Industry. *Academy of Management Journal*, 52(2), 336–354.
- Reitzig, M. & Puranam, P. (2009). Value Appropriation as an Organizational Capability: The Case of IP Protection Through Patents. *Strategic Management Journal*, 30(7), 765–789.
- Richard, O. C., Murthi, B. P. S. & Ismail, K. (2007). The Impact of Racial Diversity on Intermediate and Long-Term Performance: The Moderating Role of Environmental Context. *Strategic Management Journal*, 28(12), 1213–1233.
- Rothaermel, F. T., Hitt, M. A. & Jobe, L. A. (2006). Balancing Vertical Integration and Strategic Outsourcing: Effects on Product Portfolio, Product Success, and Firm Performance. *Strategic Management Journal*, 27(11), 1033–1056.

- Sampson, R. C. (2007). R&D Alliances and Firm Performance: The Impact of Technological Diversity and Alliance Organisation on Innovation. *Academy of Management Journal*, 50(2), 364–386.
- Sanders, W. Gerard & Hambrick, D. C. (2007). Swinging for the Fences: The Effects of CEO Stock Options on Company Risk Taking and Performance. *Academy of Management Journal*, 50(5), 1055–1078.
- Sanders, W. Gerard & Tuschke, A. (2007). The Adaption of Institutionally Contested Organizational Practices: The Emergence of Stock Option Pay in Germany. *Academy of Management Journal*, 50(1), 33–56.
- Schwab, A. & Miner, A. S. (2008). Learning in Hybride-Project Systems: The Effects of Project Performance on Repeated Collaboration. *Academy of Management Journal*, 51(6), 1117–1149.
- Semadeni, M., Canella Jr., A. A., Fraser, D. R. & Lee, D. Scott (2008). Fight or Flight: Managing Stigma in Executive Careers. *Strategic Management Journal*, 29(5), 557–567.
- Shah, R. H. & Swaminathan, V. (2008). Factors Influencing Partner Selection in Strategic Alliances: the Moderating Role of Alliance Context. *Strategic Management Journal*, 29(5), 471–494.
- Sharfman, M. P. & Fernando, C. S. (2008). Environmental Risk Management and the Cost of Capital. *Strategic Management Journal*, 29(6), 569–592.
- Shervani, T. A., Frazier, G. & Challagalla, G. (2007). The Moderating Influence of Firm Market Power on the Transaction Cost Economics Model: An Empirical Test in Forward Channel Integration Context. *Strategic Management Journal*, 28(6), 635–652.
- Shimizu, K. (2007). Prospect Theory, Behavioral Theory, and the Threat-Rigidity- Thesis: Combinative Effects on Organizational Decisions to Divest Formerly Acquired Units. *Academy of Management Journal*, 50(6), 1495–1514.
- Short, J., Ketchen, D. J., Palmer, T. B. & Hult, G. Thomas M. (2007). Firm, Strategic Group, and Industry Influences on Performance. *Strategic Management Journal*, 28(2), 147–167.
- Siebert, W. Stanley & Zubanov, N. (2009). Searching for the Optimal Level of Employee Turnover: A Study of a Large U.K. Retail Organization. *Academy of Management Journal*, 52(2), 294–313.
- Simsek, Z. (2007). CEO Tenure and Organizational Performance: An Intervening Model. *Strategic Management Journal*, 28(6), 653–662.
- Sinha, R. K. & Noble, C. H. (2008). The Adoption of Radical Manufacturing Technologies and Firm Survival. *Strategic Management Journal*, 29(9), 943–962.
- Sirmon, D. G., Gove, S. & Hitt, M. A. (2008). Resource Management in Dyadic Competitive Rivalry: The Effects of Resource Bundling and Deployment. *Academy of Management Journal*, 51(5), 919–935.
- Slater, S. F., Olson, E. M. & Hult, G. Thomas M. (2006). The Moderating Influence of Strategic Orientation on the Strategy Formation Capability-Performance Relationship. *Strategic Management Journal*, 27(12), 1221–1231.
- Somaya, D., Williamson, I. O. & Lorinkova, N. (2008). Gone but not Lost: The Different Performance Impacts of Employee Mobility Between Cooperators versus Competitors. *Academy of Management Journal*, 51(5), 936–953.
- Sorenson, O., McEvily, S., Rongrong Ren, C. & Roy, R. (2006). Niche Width Revisited: Organizational Scope, Behavior and Performance. *Strategic Management Journal*, 27(10), 915–936.

- Srivastava, A., Bartol, K. M. & Locke, E. A. (2006). Empowering Leadership in Management Teams: Effects on Knowledge Sharing, Efficacy, and Performance. *Academy of Management Journal*, 49(6), 1239–1251.
- Stam, W. & Elfring, T. (2008). Entrepreneurial Orientation and New Venture Performance: The Moderating Role of Intra- and Extraindustry Social Capital. *Academy of Management Journal*, 51(1), 91–111.
- Sun, L.-Y., Aryee, S. & Law, K. S. (2007). High-Performance Human Resource Practices, Citizenship Behaviour, and Organizational Performance: A Relational Perspective. *Academy of Management Journal*, 50(3), 558–577.
- Szulanski, G. & Jensen, R. J. (2006). Presumptive Adaption and the Effectiveness of Knowledge Transfer. *Strategic Management Journal*, 27(10), 937–957.
- Takeuchi, R., Shay, J. P. & Li, J. (2008). When Does Decision Autonomy Increase Expatriate Managers' Adjustment? An Empirical Test. *Academy of Management Journal*, 51(1), 45–60.
- Tanriverdi, H. & Lee, C.-H. (2008). Within-Industry Diversification and Firm Performance in the Presence of Network Externalities: Evidence from the Software Industry. *Academy of Management Journal*, 51(2), 381–397.
- Terlaak, A. & King, A. A. (2007). Follow the Small? Information-Revealing Adoption Bandwagons when Observers Expect Larger Firms to Benefit More from Adoption. *Strategic Management Journal*, 28(12), 1167–1185.
- Thornhill, S. & White, R. E. (2007). Strategic Purity: A Multi-Industry-Evaluation of Pure Vs. Hybrid Business Strategies. *Strategic Management Journal*, 28(5), 553–561.
- Tiwana, A. (2008). Does Technological Modularity Substitute for Control? A Study of Alliance Performance in Software Outsourcing. *Strategic Management Journal*, 29(7), 769–780.
- Tiwana, A. & Keil, M. (2007). Does Peripheral Knowledge Complement Control? An Empirical Test in Technology Outsourcing Alliances. *Strategic Management Journal*, 28(6), 623–634.
- Tong, T. W., Reuer, J. J. & Peng, M. W. (2008). International Joint Ventures and the Value of Growth Options. *Academy of Management Journal*, 51(5), 1014–1029.
- Tsang, E. W. K. & Yip, P. S. L. (2007). Economic Distance and the Survival of Foreign Direct Investments. *Academy of Management Journal*, 50(5), 1156–1168.
- Uhlenbruck, K., Hitt, M. A. & Semadeni, M. (2006). Market Value Effects of Acquisitions Involving Internet Firms: A Resource-based Analysis. *Strategic Management Journal*, 27(10), 899–913.
- Uotila, J., Maula, M., Keil, T. & Zahra, S. A. (2009). Exploration, Exploitation, and Financial Performance: Analysis of S&P 500 Corporations. *Strategic Management Journal*, 30(2), 221–231.
- Voss, G. B., Sirdeshmukh, D. & Voss, Z. Giraud (2008). The Effects of Slack Resources and Environmental Threat on Product Exploration and Exploitation. *Academy of Management Journal*, 51(1), 147–164.
- Vroom, G. & Gimeno, J. (2007). Ownership Form, Managerial Incentives, and the Intensity of Rivalry. *Academy of Management Journal*, 50(4), 901–922.
- Wade, J. B., Porac, J. F., Pollock, T. G. & Graffin, S. D. (2006). The Burden of Celebrity: The Impact of CEO Certification Contests on CEO Pay and Performance. *Academy of Management Journal*, 49(4), 643–660.
- Wadhwa, A. & Kotha, S. (2006). Knowledge Creation Through External Venturing: Evidence from the Telecommunications Equipment Manufacturing Industry. *Academy of Management Journal*, 49(4), 819–835.

- Wan, W. P. & Yiu, D. W. (2009). From Crisis to Opportunity: Environmental Jolt, Corporate Acquisitions, and Firm Performance. *Strategic Management Journal*, 30(7), 791–801.
- Wang, L. & Zajac, E. J. (2007). Alliance or Acquisition? A Dyadic Perspective on Interfirm Resource Combinations. *Strategic Management Journal*, 28(13), 1291–1317.
- Wasserman, N. (2006). Stewards, Agents, and the Founder Discount: Executive Compensation in New Ventures. *Academy of Management Journal*, 49(5), 960–976.
- Weigelt, C. (2009). The Impact of Outsourcing New Technologies on Integrative Capabilities and Performance. *Strategic Management Journal*, 30(6), 595–616.
- Weigelt, C. & Sarkar, M. B. (2009). Learning from Supply-Side Agents: The Impact of Technology Solution Providers' Experiential Diversity on Clients' Innovation Adoption. *Academy of Management Journal*, 52(1), 37–60.
- Westphal, J. D. & Stern, I. (2007). Flattery Will Get You Everywhere (Especially If You Are a Male Caucasian): How Integration, Boardroom Behaviour, and Demographic Minority Status Affect Additional Board Appointments at U.S. Companies. *Academy of Management Journal*, 50(2), 267–288.
- Wiersema, M. F. & Bowen, H. P. (2008). Corporate Diversification: The Impact of Foreign Competition, Industry Globalization, and Product Diversification. *Strategic Management Journal*, 29(2), 115–132.
- Williams, C. (2007). Transfer in Context: Replication and Adaption in Knowledge Transfer Relationships. *Strategic Management Journal*, 28(9), 867–889.
- Zatzick, C. D. & Iverson, R. (2006). High-Involvement Management and Workforce Reduction: Competitive Advantage or Disadvantage? *Academy of Management Journal*, 49(5), 999–1015.
- Zhang, X., Bartol, K. M., Smith, K. G., Pfarrer, M. D. & Khanin, D. M. (2008). CEOs on the Edge: Earnings Manipulation and Stock-Based Incentive Misalignment. *Academy of Management Journal*, 51(2), 241–258.
- Zhang, Y. (2007). Information Asymmetry and the Dismissal of Newly Appointed CEOs: An Empirical Investigation. *Strategic Management Journal*, 28(8), 859–872.
- Zhang, Y. & Wiersema, M. F. (2009). Stock Market Reaktion to CEO Certification: The Signaling Role of CEO Background. *Strategic Management Journal*, 30(7), 693–710.
- Zhou, K. Zheng, Li, J. Juan, Zhou, N. & Su, C. (2008). Market Orientation, Job Satisfaction, Product Quality, and Firm Performance: Evidence from China. *Strategic Management Journal*, 29(9), 985–1000.
- Zott, C. & Amit, R. (2008). The Fit Between Product Market Strategy and Business Modell: Implications for Firm Performance. *Strategic Management Journal*, 29(1), 1–26.

Literaturverzeichnis III: Verzeichnis der in die Metaanalyse einbezogenen Primärstudien

- Bae, J., Chen, S.-j., Wan, T. Wan David, Lawler, J. J. & Walumbwa, F. Ochieng (2003). Human Resource Strategy and Firm Performance in Pacific Rim Countries. *International Journal of Human Resource Management*, 14(8), 1308–1332.
- Bae, J. & Lawler, J. J. (2000). Organizational and HRM-Strategies in Korea: Impact on Firm Performance in an Emerging Economy. *Academy of Management Journal*, 42(3), 502–517.
- Bae, J. & Gargiulo, M. (2004). Partner Substitutability, Alliance Network Structure and Firm Profitability in the Telecommunications Industry. *Academy of Management Journal*, 47(6), 843–859.
- Barua, A., Konana, P., Whinston, A. & Yin, F. (2004). An Empirical Investigation of Net-Enabled Business Value. *MIS Quarterly*, 28(4), 585–620.
- Batjargal, B. (2005). Entrepreneurial Versatility, Resources and Firm Performance in Russia: A Panel Study. *Journal of Entrepreneurship and Innovation Management*, 5(3), 284–297.
- Bergh, D. D. (2001). Executive Retention and Acquisition Performance: A Test of Opposing Views on the Influence of Organizational Tenure. *Journal of Management*, 27(5), 603–622.
- Berman, S. L., Down, J. & Hill, C. W. I. (2002). Tacit Knowledge as a Source of Competitive Advantage in the National Basketball Association. *Academy of Management Journal*, 45(1), 13–31.
- Capron, L. (1999). The Long-Term Performance of Horizontal Acquisitions. *Strategic Management Journal*, 20(11), 987–1018.
- Carmeli, A. & Tishler, A. (2004). The Relationships between Intangible Organizational Elements and Organizational Performance. *Strategic Management Journal*, 25(13), 1257–1278.
- Carolis, D. Marie de (2003). Competences and Imitability in the Pharmaceutical Industry: An Analysis of Their Relationship with Firm Performance. *Journal of Management*, 29(1), 27–50.
- Carow, K., Heron, R. & Saxton, T. (2004). Do Early Birds Get the Returns? An Empirical Investigation of Early Mover Advantages in Acquisitions. *Strategic Management Journal*, 25(6), 563–585.
- Carpenter, M. A., Sanders, W. Gerard & Gregersen, H. B. (2001). Bundling Human Capital With Organizational Context: The Impact of International Assignment Experience on Multinational Firm Performance and CEO Pay. *Academy of Management Journal*, 44(3), 493–511.
- Chadwick, C., Hunter, L. W. & Walston, S. L. (2004). Effects of Downsizing Practices on the Performance of Hospitals. *Strategic Management Journal*, 25(5), 405–427.
- Challis, D., Samson, D. & Lawson, B. (2002). Integrated Manufacturing, Employee and Business Performance: Australian and New Zealand Evidence. *International Journal of Production Research*, 40(8), 1941–1964.
- Chan, L. L. M., Shaffer, M. A. & Snape, E. (2004). In Search of Sustained Competitive Advantage: The Impact of Organizational Culture, Competitive Strategy and Human Resource Management Practices on Firm Performance. *International Journal of Human Resource Management*, 15(1), 17–35.
- Chandler, G. N. & Hanks, S. H. (1994). Market Attractiveness, Resource-Based Capabilities, Venture Strategies, and Venture Performance. *Journal of Business Venturing*, 9(4), 331–349.
- Cho, H.-J. & Pucik, V. (2005). Relationship Between Innovativeness, Quality, Growth, Profitability, and Market Value. *Strategic Management Journal*, 26(6), 555–575.

- Christmann, P. (2000). Effects of "Best Practices" of Environmental Management on Cost Advantage: The Role of Complementary Assets. *Academy of Management Journal*, 43(4), 663–680.
- Collins, C. J. & Clark, K. D. (2003). Strategic Human Resource Practices, Top Management Team Social Networks, and Firm Performance: The Role of Human Resource Practices in Creating Organizational Competitive Advantage. *Academy of Management Journal*, 46(6), 740–751.
- Combs, J. G. & Ketchen, D. J. (1999). Explaining Interfirm Cooperation and Performance: Toward a Reconciliation of Predictions from the Resource-based View and Organizational Economics. *Strategic Management Journal*, 20(9), 867–888.
- Datta, D. K., Guthrie, J. P. & Wright, P. M. (2005). Human Resource Management and Labor Productivity: Does Industry Matter? *Academy of Management Journal*, 48(1), 135–145.
- Deephouse, D. L. (2000). Media Reputation as a Strategic Resource: An Integration of Mass Communication and Resource-Based Theories. *Journal of Management*, 26(6), 1091–1112.
- Dhanaraj, C. & Beamish, P. W. (2003). A Resource-based Approach to the Study of Export Performance. *Journal of Small Business Management*, 41(3), 242–261.
- Douglas, T. J. & Ryman, J. A. (2003). Understanding Competitive Advantage in the General Hospital Industry: Evaluating Strategic Competencies. *Strategic Management Journal*, 24(4), 333–347.
- Dröge, C., Claycomb, C. & Germain, R. (2003). Does Knowledge Mediate the Effect of Context on Performance? Some Initial Evidence. *Decision Sciences*, 34(3), 541–568.
- Edelman, L. F., Brush, C. G. & Manolova, T. (2005). Co-Alignment in the Resource-Performance Relationship: Strategy as Mediator. *Journal of Business Venturing*, 20(3), 359–383.
- Fey, C. F., Björkman, I. & Pavlovskaya Antonina (2000). The Effect of Human Resource Management Practices on Firm Performance in Russia. *International Journal of Human Resource Management*, 11(1), 1–18.
- Gong, Y. (2003). Subsidiary Staffing in Multinational Enterprises: Agency, Resources, and Performance. *Academy of Management Journal*, 46(6), 728–739.
- Guerrero, S. & Barraud-Didier, V. (2004). High-Involvement Practices and Performance of French Firms. *International Journal of Human Resource Management*, 15(8), 1408–1423.
- Guest, D. E., Michie, J., Conway, N. & Sheehan, M. (2003). Human Resource Management and Corporate Performance in the UK. *British Journal of Industrial Relations*, 41(2), 291–314.
- Harel, G. H. & Tzafrir, S. S. (1999). The Effect of Human Resource Management Practices on the Perceptions of Organizational and Market Performance of the Firm. *Human Resource Management*, 38(3), 185–200.
- Hatch, N. & Dyer, J. H. (2004). Human Capital and Learning as a Source of Sustainable Competitive Advantage. *Strategic Management Journal*, 25(12), 1155–1178.
- Hayton, J. C. (2003). Strategic Human Capital Management in SMEs: An Empirical Study of Entrepreneurial Performance. *Human Resource Management*, 42(4), 375–391.
- Hayton, J. C. (2005). Competing in the New Economy: The Effect of Intellectual Capital on Corporate Entrepreneurship in High-Technology New Ventures. *R&D Management*, 35(2), 137–155.
- Hillman, A. J. & Keim, G. D. (2001). Shareholder Value, Stakeholder Management, and Social Issues: What's the Bottom Line? *Strategic Management Journal*, 22(2), 125–139.
- Hitt, M. A., Bierman, L., Shimizu, K. & Kochhar, R. (2001). Direct and Moderating Effects of Human Capital on Strategy and Performance in Professional Service Firms: A Resource-Based Perspective. *Academy of Management Journal*, 44(1), 13–28.

- Hitt, M. A., Bierman, L., Uhlenbruck, K. & Shimizu, K. (2006). The Importance of Resources in the Internationalization of Professional Service Firms: the Good, the Bad, and the Ugly. *Academy of Management Journal*, 49(6), 1137–1157.
- Hitt, M. A., Hoskisson, R. E. & Kim, H. (1997). International Diversifikation: Effects of Innovation and Firm Performance in Product-Diversified Firms. *Academy of Management Journal*, 40(4), 767–798.
- Hooley, G. J., Greenley, G. E., Cadogan, J. W. & Fahy, J. (2005). The Performance Impact of Marketing Resources. *Journal of Business Research*, 58(1), 18–27.
- Hult, G. Thomas M. & Ketchen, D. J. (2001). Does Market Orientation Matter? A Test of the Relationship Between Positional Advantage and Performance. *Strategic Management Journal*, 22(9), 899–906.
- Huselid, M. A. (1995). The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. *Academy of Management Journal*, 38(3), 635–672.
- Huselid, M. A., Jackson, S. E. & Schuler, R. S. (1997). Technical and Strategic Human Resource Management Effectiveness as Determinants of Firm Performance. *Academy of Management Journal*, 40(1), 171–188.
- Irwin, J. G., Hoffman, J. J. & Lamont, B. T. (1998). The Effect of the Acquisition of Technological Innovations on Organizational Performance: A Resource-based View. *Journal of Engineering and Technology Management*, 15(1), 25–54.
- Judge, W. Q. & Douglas, T. J. (1998). Performance Implications of Incorporating Natural Environmental Issues into the Strategic Planning Process: An Empirical Assessment. *Journal of Management Studies*, 35(2), 241–262.
- Judge, W. Q. & Elenkov, D. (2005). Organizational Capacity for Change and Environmental Performance: An Empirical Assessment of Bulgarian Firms. *Journal of Business Research*, 58(7), 893–901.
- Keller, R. T. (2004). A Resource-Based Study of New Product Development: Predicting Five-Year Later Commercial Success and Speed to Market. *International Journal of Innovation Management*, 8(3), 243–260.
- King, A. Wilcox & Zeithaml, C. P. (2001). Competences and Firm Performance: Examining the Causal Ambiguity Paradox. *Strategic Management Journal*, 22(1), 75–99.
- Klassen, R. D. & Whybark, D. Clay (1999). The Impact of Environmental Technologies on Manufacturing Performance. *Academy of Management Journal*, 42(6), 599–615.
- Kor, Y. Y. & Mahoney, J. T. (2005). How Dynamics, Management, and Governance of Resource Deployments Influence Firm-Level Performance. *Strategic Management Journal*, 26(5), 489–496.
- Lee, C., Lee, K. & Pennings, J. M. (2001). Internal Capabilities, External Networks, and Performance: A Study on Technology-Based Ventures. *Strategic Management Journal*, 22(6/7), 615–640.
- Lee, J. & Miller, D. (1999). People Matter: Commitment to Employees, Strategy and Performance in Korean Firms. *Strategic Management Journal*, 20(6), 579–593.
- Lee, R. P. & Grewal, R. (2004). Strategic Responses to New Technologies and Their Impact on Firm Performance. *Journal of Marketing*, 68(4), 157–171.
- Lerner, M. & Almor, T. (2002). Relationships among Strategic Capabilities and the Performance of Women-Owned Small Ventures. *Journal of Small Business Management*, 40(2), 109–125.
- Li, J., Lam, K. & Qian, G. (2001). Does Culture Affect Behavior and Performance of Firms? The Case of Joint Ventures in China. *Journal of International Business Studies*, 32(1), 115–131.

- Luo, X., Griffith, D. A., Liu, S. S. & Shi, Y.-Z. (2004). The Effects of Customer Relationships and Social Capital on Firm Performance: A Chinese Business Illustration. *Journal of International Marketing*, 12(4), 25–45.
- Menguc, B., Auh, S. & Shih, E. (2007). Transformational Leadership and Market Orientation: Implications for the Implementation of Competitive Strategies and Business Unit Performance. *Journal of Business Research*, 60(4), 314–321.
- Menguc, B. & Barker, T. (2005). Re-examining Field Sales Unit Performance: Insights From the Resource-Based View and Dynamic Capabilities Perspectives. *European Journal of Marketing*, 39(7/8), 885–909.
- Menguc, B. & Ozanne, L. K. (2005). Challenges of the "Green Imperative": A Natural Resource-Based Approach of the Environmental Orientation-Business Performance Relationship. *Journal of Business Research*, 58(4), 430–438.
- Miller, D. & Shamsie, J. (1996). The Resource-Based View of the Firm in Two Environments: The Hollywood Film Studios from 1936 to 1965. *Academy of Management Journal*, 39(3), 519–543.
- Miller, D. J. (2006). Technological Diversity, Related Diversifikation, and Firm Performance. *Strategic Management Journal*, 27(7), 601–619.
- Morrow Jr., J. L., Sirmon, D. G., Hitt, M. A. & Holcomb, T. R. (2007). Creating Value in the Face of Declining Performance: Firm Strategies and Organizational Recovery. *Strategic Management Journal*, 28(3), 271–283.
- O'Shaugnessy, K. C., Gedajlovic, E. & Reinmoeller, P. (2007). The Influence of Firm, Industry and Network on the Corporate Social Performance of Japanese Firms. *Asia Pacific Journal of Management*, 24(3), 283–303.
- Park, H. Jeong, Mitsunashi, H., Fey & Björkmann, I. (2003). The Effect of Human Resource Management Practices on Japanese MNC Subsidiary Performance: A Partial Mediating Model. *International Journal of Human Resource Management*, 14(8), 1391–1406.
- Park, S. Ho & Luo, Y. (2001). Guanxi and Organizational Dynamics: Organizational Networking in Chinese Firms. *Strategic Management Journal*, 22(5), 455–477.
- Patterson, M. G., West, M. A. & Wall, T. D. (2004). Integrated Manufacturing, Empowerment, And Company Performance. *Journal of Organizational Behavior*, 25(5), 641–665.
- Peng, M. W. & York, A. S. (2001). Behind Intermediary Performance in Export Trade: Transactions, Agents, and Resources. *Journal of International Business Studies*, 32(2), 327–346.
- Perry-Smith, J. E. & Blum, T. C. (2000). Work-Family Human Resource Bundles and Perceived Organizational Performance. *Academy of Management Journal*, 43(6), 1107–1117.
- Phua, F. T. T. (2006). Predicting Construction Firm Performance: An Empirical Assessment of the Differential Impact Between Industry and Firm-Specific Factors. *Construction Management and Economics*, 24(3), 309–320.
- Powell, T. C. (1992). Organizational Alignment as Competitive Advantage. *Strategic Management Journal*, 13(2), 119–134.
- Powell, T. C. & Dent-Micallef, A. (1997). Information Technology as Competitive Advantage: The Role of Human, Business, and Technology Resources. *Strategic Management Journal*, 18(5), 375–405.
- Richard, O. C. (2000). Racial Diversity, Business Strategy, and Firm Performance: A Resource-Based View. *Academy of Management Journal*, 43(2), 164–177.
- Richard, O. C., McMillan, A., Chadwick, K. & Dwyer, S. (2003). Employing an Innovation Strategy in Racially Diverse Workforces: Effects on Firm Performance. *Group and Organization Management*, 28(1), 107–126.

- Roberts, P. W. & Dowling, G. R. (2002). Corporate Reputation and Sustained Superior Financial Performance. *Strategic Management Journal*, 23(12), 1077–1093.
- Robins, J. A. & Wiersema, M. F. (1995). A Resource-based Approach to the Multibusiness Firm: Empirical Analysis of Portfolio Interrelationships and Corporate Financial Performance. *Strategic Management Journal*, 16(4), 277–299.
- Roth, K. (1995). Managing International Interdependence: CEO Characteristics in a Resource-Based Framework. *Academy of Management Journal*, 38(1), 200–231.
- Rothaermel, F. T. & Thursby, M. (2005). University-Incubator Firm Knowledge Flows: Assessing their Impact on Incubator Firm Performance. *Research Policy*, 34(3), 305–320.
- Russo, M. V. & Fouts, P. A. (1997). A Resource-Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal*, 40(3), 534–559.
- Shaw, J. D., Gupta, N. & Delery, J. E. (2005). Alternative Conceptualizations of the Relationship Between Voluntary Turnover and Organizational Performance. *Academy of Management Journal*, 48(1), 50–68.
- Sher, P. J. & Yang, P. Y. (2005). The Effects of Innovative Capabilities and R&D Clustering on Firm Performance: The Evidence of Taiwans Semiconductor Industry. *Technovation*, 25(1), 33–43.
- Short, J., Palmer, T. B. & Ketchen, D. J. (2002). Resource-Based and Strategic Group Influences on Hospital Performance. *Health Care Management Review*, 27(4), 7–17.
- Simonin, B. L. (1997). The Importance of Collaborative Know-How: An Empirical Test of the Learning Organization. *Academy of Management Journal*, 40(5), 1150–1174.
- Skaggs, B. C. & Youndt, M. (2004). Strategic Positioning, Human Capital, and Performance in Service Organisations: A Customer Interaction Approach. *Strategic Management Journal*, 25(1), 85–99.
- Slootegraaf, R. J., Moorman, C. & Inman, J. Jeffrey (2003). The Role of Firm Resources in Returns to Market Deployment. *Journal of Marketing Research*, 40(3), 295–309.
- Song, M., Di Benedetto, A. & Nason, R. W. (2007). Capabilities and Financial Performance: The Moderating Effect of Strategic Type. *Journal of the Academy of Marketing*, 35(1), 18–34.
- Spannos, Y. E. & Lioukas, S. (2001). An Examination Into the Causal Logic of Rent Generation: Contrasting Porter's Competitive Strategy Framework and the Resource-based Perspective. *Strategic Management Journal*, 22(10), 907–934.
- Steensma, H. Kevin & Corley, K. G. (2000). On the Performance of Technology-Sourcing Partnerships: The Interaction Between Partner Interdependence and Technology Attributes. *Academy of Management Journal*, 43(6), 1045–1067.
- Sun, L.-Y., Aryee, S. & Law, K. S. (2007). High-Performance Human Resource Practices, Citizenship Behaviour, and Organizational Performance: A Relational Perspective. *Academy of Management Journal*, 50(3), 558–577.
- Tanriverdi, H. & Venkatraman, N. (2005). Knowledge Relatedness and the Performance of Multibusiness Firms. *Strategic Management Journal*, 26(2), 97–119.
- Thompson, M. & Heron, P. (2005). Management Capability and High Performance Work Organization. *International Journal of Human Resource Management*, 16(6), 1029–1048.
- Tippins, M. J. & Sohi, R. S. (2003). IT Competency and Firm Performance: Is Organizational Learning a Missing Link? *Strategic Management Journal*, 24(8), 745–761.
- Wang, C. K. & Ang, B. Lian (2004). Determinants of Venture Performance in Singapore. *Journal of Small Business Management*, 42(4), 347–363.

- Wang, Y., Lo, H.-P. & Yang, Y. (2004). The Constituents of Core Competencies and Firm Performance: Evidence from High-Technology Firms in China. *Journal of Engineering and Technology Management*, 21(4), 249–280.
- Wiklund, J. & Shepherd, D. (2003). Knowledge-Based Resources, Entrepreneurial Orientation, and the Performance of Small and Medium-Sized Businesses. *Strategic Management Journal*, 24(13), 1307–1314.
- Wright, P. M., Gardner, T. M., Moynihan, L. M. & Allen, M. R. (2005). The Relationship Between HR Practices and Firm Performance: Examining Causal Order. *Personnel Psychology*, 58(2), 409–446.
- Wright, P. M., McCormick, B., Sherman, W. Scott & McMahan, G. C. (1999). The Role of Human Resource Practices in Petro-Chemical Refinery Performance. *International Journal of Human Resource Management*, 10(4), 551–571.
- Yeoh, P.-L. (2004). International Learning: Antecedents and Performance Implications Among Newly Internationalizing Companies in an Exporting Context. *International Marketing Review*, 21(4/5), 511–535.
- Yip, G. S., Biscarri, J. Gomez & Monti, J. A. (2000). The Role of the Internationalization Process in the Performance of Newly Internationalizing Firms. *Journal of International Marketing*, 8(3), 10–35.
- Young, G., Sapienza, H. & Baumer, D. (2003). The Influence of Flexibility in Buyer-Seller-Relationships on the Productivity of Knowledge. *Journal of Business Research*, 56(6), 443–451.
- Zahra, S. A., Ireland, R. Duane & Hitt, M. A. (2000). International Expansion by New Venture Firms: International Diversity, Mode of Market Entry, Technological Learning, and Performance. *Academy of Management Journal*, 43(5), 925–950.
- Zahra, S. A. & Nielsen, A. P. (2002). Sources of Capabilities, Integration and Technology Commercialization. *Strategic Management Journal*, 23(5), 377–398.

A.2 Tabellen

Tabelle A.1: Vollständige Daten Literaturreview

Tabelle A.1 ist beiliegendem Datenträger zu entnehmen. Dateiname: Daten_Literaturreview.xlsx).

Tabelle A.2: Vollständige Daten Metaanalyse

Auch dieser Datensatz findet sich auf beiliegendem Datenträger, unter dem Dateinamen Daten_Metaanalyse.xlsx.

Tabelle A.3: Die verwendeten Performanceoperationalisierungen in der qualitativen Literaturreview

	N	Prozent	Prozent der Fälle
Altmans Z	1	0,4	0,5
Gewinn	2	0,8	1,0
ROA, ROI, ROE, ROCE	57	21,5	28,6
ROS	10	3,8	5,0
CAR	16	6,0	8,0
Jensens Alpha	2	0,8	1,0
Marktkapitalisierung	5	1,9	2,5
Price-Earnings-Ratio	1	0,4	0,5
Market-to-Book Ratio	22	8,3	11,1
Total Shareholder Return	12	4,5	6,0
Kapitalkosten	2	0,8	1,0
Asset Turnover	1	0,4	0,5
Produzierter Ausschuss	1	0,4	0,5
Fluktuation	1	0,4	0,5
Innovationsoutput	13	4,9	6,5
Pünktlichkeit Auslieferung	2	0,8	1,0
Kapazitätsauslastung	3	1,1	1,5
Kostenfunktion	1	0,4	0,5
Lagerumschlag	2	0,8	1,0
Marktanteil	4	1,5	2,0
Mitarbeiterzahl	2	0,8	1,0
Umsatz/Mitarbeiter	8	7,0	4,0
Überleben/Ausscheiden	15	5,7	7,5
Umsatz	21	7,9	10,6
Externe Experten	4	1,5	2,0
Interne Experten	33	12,5	16,6
Sonstige	24	9,1	12,1
Gesamt	265	100	133,2

Insgesamt betrachtete Artikel mit Leistungsbezug: 199

Tabelle A.4: Herkunft der in die Metaanalyse einbezogenen Primärstudien nach Journal

Journal	Häufigkeit	Prozent
AMJ	37	25,9
Asia Pacific J o Management	2	1,4
British J of Industrial Relations	3	2,1
Construction Management	1	,7
Decision Science	1	,7
European J of Marketing	1	,7
Group and Organization	1	,7
Health Care Management	1	,7
Human Resource Management	2	1,4
Internat J o HRM	1	,7
Internat J o Innovation	1	,7
Internat J of HRM	6	4,2
Internat J of Production	1	,7
Internat Marketing Review	1	,7
J o Business Research	7	4,9
J o Business Venturing	2	1,4
J o Engineering a Tech	2	1,4
J o Entrepreneurship a	1	,7
J o Internat Business	4	2,8
J o Internat Marketing	3	2,1
J o Management	5	3,5
J o Marketing	1	,7
J o Marketing Research	1	,7
J o Organizational Behaviour	2	1,4
J o Small Business Man	5	3,5
J o t Academy of Marketing	1	,7
J o Management Studies	1	,7
MIS Quarterly	1	,7
Personell Psychology	1	,7
R&D Management	1	,7
Research Policy	2	1,4
SMJ	42	29,4
Technovation	2	1,4
Gesamt	143	100,0

Tabelle A.5: Deskriptive Statistik Korrekturfaktoren Metaanalyse

	N Gültig	N Feh- lend	Gesamt	Mittelwert	Varianz	Max	Min
Reliabilitätskoeffizient Ressource	28	115	143	0,7868	0,003	0,91	0,67
Reliabilitätskoeffizient Performance	18	125	143	0,7967	0,008	0,94	0,59

A.3 Inhaltlicher Anhang

A.3.1 Regionaler Fokus der Strategischen Managementforschung

Im Rahmen der Diskussion der Populationsselektion in Abschnitt 3.2 fiel auf, dass die Betrachtung von Samples aus dem US-amerikanischen Bereich als Referenzwert angesehen wird, wohingegen Abweichungen von diesem Normalfall in der Regel ausführlich begründet werden. Betrachtungen US-amerikanischer Unternehmen stellen den Normalfall in der Strategischen Managementforschung dar, wie folgende Tabelle zeigt. Die Studien der qualitativen Literaturreview wurden danach kodiert, aus welcher Region das jeweilige Sample gezogen wurde.

Tabelle A.6: Populationseingrenzung: Land

	Häufigkeit	Prozent
Nordamerika	124	72,9 %
Europa	18	10,6 %
Asien	23	13,5 %
Afrika	3	1,8 %
Südamerika	1	0,6 %
Australien und Ozeanien	1	0,6 %
Gesamt	170	100 %

N = 170, fehlend: 19¹³⁹

Die Vermutung liegt nahe, dass diese Zentrierung auf die Vereinigten Staaten mit einer entsprechenden Dominanz US-amerikanischer Autoren im Diskursraum der Strategischen Managementforschung zu erklären ist, wie auch etwa von Hitt et al. (2004) gemutmaßt wurde. Wie ich in Kapitel 4 zeigte, stellen die in der qualitativen Literaturreview untersuchten Journals wesentliche Kommunikationskanäle für den Diskursraum der Strategischen Managementforschung im internationalen Kontext dar (vergleiche Schulze, Warning & Wiermann, 2008). Die Sprache dieses Diskursraumes ist Englisch, entsprechend finden sich keine nicht-englischsprachigen einschlägigen Zeitschriften im Bereich „A“ und „A+“. Es ist daher davon auszugehen, dass Wissenschaftler, die qualitativ hochwertig im Bereich der Strategischen Managementforschung auf internationalem Niveau arbeiten, unabhängig von ihrer eigenen Herkunft danach streben, in diesen Journals zu publizieren. Entsprechend ist von einer Auswertung dieser Journals auch ein tatsächlicher Überblick über den internationalen Status Quo der Strategischen Managementforschung, auch in Hinsicht auf die relevanten wissenschaftlichen Akteure, zu erwarten.

Um diese Idee zu überprüfen kodiere ich die Beiträge der qualitativen Literaturreview nach nationaler Herkunft der Autoren.¹⁴⁰ Ich ordne Beiträge einem bestimmten Herkunftsraum zu, wenn die Mehrzahl der Autoren dieses Beitrags beruflich in diesem Land tätig ist. Sowohl im Strategic Management Journal als auch im Academy of Management Journal ist der Arbeitgeber des jeweiligen Autors angeführt, mittels Internet-Recherche bin ich so in der Lage, diese Arbeitgeber geographisch einzuordnen. Da keine der betrachteten Zeitschriften die Namen der Autoren sortiert (etwa alphabetisch) und mit der Erstnennung ein größerer Prestigege-

¹³⁹ Fehlende Werte bedeuten hier, dass entweder keine oder keine nachvollziehbare regionale Eingrenzung des Samples erfolgte, es sich um Fallstudien oder Meta-Analysen handelt.

¹⁴⁰ Ein solches Vorgehen, die Einbeziehung von Faktoren aus dem Entstehungsumfeld von Studien, empfiehlt etwa Lipsey (1994).

winn einhergeht als mit einer Nennung an folgender Stelle gehe ich im Weiteren davon aus, dass die erstgenannten Autoren im Zweifel den größten Einfluss auf die erstellten Artikel hatten. Entsprechend ordne ich Beiträge, die nach obigem Algorithmus keine eindeutige Zuordnung ergeben haben, dem Herkunftsland des Arbeitgebers des erstgenannten Autors zu.

Sonderfälle: Bei Institutionen, die über Niederlassungen in verschiedenen Ländern verfügen, ordne ich nach der Ursprungsinstitution zu. INSEAD beispielsweise verfügt über bedeutende Niederlassungen in Frankreich und Singapur, der Campus in Singapur wurde jedoch erst im Jahr 2000 eingeweiht, entsprechend betrachte ich Frankreich als Heimatland dieser Institution und ordne alle Beiträge von Autoren, die bei INSEAD beschäftigt sind, Frankreich zu. Sind Autoren an mehreren Institutionen gleichzeitig beschäftigt und findet sich nach oben beschriebenen Algorithmus keine eindeutige Zuordnung, nehme ich diese Beiträge aus der Zählung. Dies war bei einem der betrachteten 199 Artikel der Fall.

Folgende Tabelle zeigt absolute und relative Häufigkeiten der Herkunft:

Tabelle A.7: Herkunft der Autoren

Land	Häufigkeit	Prozent
USA	141	70,9
Kanada	11	5,5
UK	3	1,5
Deutschland	2	1,0
Frankreich	5	2,5
Italien	2	1,0
Spanien	3	1,5
Finnland	5	2,5
Dänemark	1	,5
Norwegen	1	,5
Niederlande	3	1,5
Schweiz	1	,5
China	2	1,0
Hongkong	5	2,5
Korea	3	1,5
Singapur	5	2,5
Australien	1	,5
Neuseeland	1	,5
Brasilien	1	,5
Vereinigte Arabische Emirate	1	,5
Türkei	1	,5
Zuordnung nicht möglich	1	,5
Gesamt	199	100,0

Offensichtlich ist eine klare Dominanz US-amerikanischer Autoren. 71 Prozent aller untersuchten Beiträge haben ihren Ursprung an US-amerikanischen Forschungseinrichtungen. Nimmt man noch Kanada, die zweithäufigste Ausprägung, hinzu lässt sich festhalten, dass mehr als Dreiviertel aller Beiträge der Strategischen Managementforschung aus dem nord-amerikanischen Raum stammen. In Europa spielen noch Frankreich und Finnland mit jeweils fünf Beiträgen (2,5 Prozent) eine gewisse Rolle, im asiatischen Raum die forschungsstarken Hochschulen in Singapur und Hongkong. In Frankreich lässt sich zudem eine besondere Dominanz des INSEAD feststellen, vier der fünf Artikel haben ihren Ursprung an dieser Business School. Es steht zu vermuten, dass sich eine ähnliche Clusterung auch in anderen Ländern feststellen und der Output an A-Journals sich damit wesentlich mit der Institution, der die Autoren angehören, erklären ließe; eine Aufteilung nach einzelnen Institutionen habe ich jedoch nicht durchgeführt, daher muss diese Detailfrage zukünftigen Forschungen überlassen bleiben. Deutschland spielt mit zwei publizierten Artikeln in der internationalen Strategischen Managementforschung keine Rolle, im AMJ findet sich im Fachgebiet der Strategischen Managementforschung kein einziger Beitrag aus Deutschland. Aus methodischen Gründen (qualitativer Ansatz) ist eine Verallgemeinerbarkeit der Aussagen auf andere Journals oder Forschungsbereich naturgemäß nicht möglich.

A.3.2 Weitere Methodenprobleme in der Strategischen Managementforschung

Über die drei im Rahmen dieser Arbeit diskutierten Problemfelder hinaus werden in der Literatur, die sich kritisch mit Methodenproblemen in der quantitativen Sozialforschung im Allgemeinen und der Strategischen Managementforschung im Besonderen auseinandersetzt, noch einige weitere potentielle Methodenprobleme angesprochen, die ich im Folgenden kurz vorstellen und in ihrer Relevanz einordnen möchte. Hierbei handelt es sich zumeist um Probleme der Messmethodik und der Konstruktoperationalisierung. Der Abschnitt beginnt mit allgemeinen Methodenproblemen, wie sie im Standardwerk von Hunter & Schmidt (2007) zusammengefasst sind, und betrachtet anschließend weitere spezifische Problemfelder der Strategischen Managementforschung.

A.3.2.1 Allgemeine Methodenprobleme in der quantitativen Sozialforschung

In ihrem Standardwerk zu methodischen Artefakten¹⁴¹ in der quantitativen Forschung erstellen Hunter & Schmidt (2007) einen umfassenden Katalog elf systematischer Probleme, der im Folgenden kurz in den Kontext dieser Arbeit eingeordnet werden soll:

1. Sampling error – zufällige Schwankungen der betrachteten Korrelation je nach Stichprobenziehung. Allgemein treten in quantitativen Studien Schwierigkeiten auf, wenn die Stichprobe schlecht gezogen wird. Der sampling error beschreibt dieses Problem. Betrachtet wird die Differenz zwischen der beobachteten Beziehung zwischen zwei Variablen und einer als existent angenommenen wahren Beziehung in der jeweiligen Grundgesamtheit. Je nachdem wie gut die gezogene Stichprobe die Gesamtpopulation abbildet, wird die tatsächliche Bezie-

¹⁴¹ Unter Artefakten verstehen die Autoren Ergebnisse, beziehungsweise Ergebnisverzerrungen, bei denen es sich um „artefactual, or man made errors and not properties of nature“ handelt, also Verzerrungen oder Ergebnisse, die durch methodische Schwierigkeiten entstehen und so die wahren Zusammenhänge verschleiern oder verfälschen (Hunter & Schmidt, 2007, 33).

hung von der beobachteten mehr oder weniger stark abweichen. Da es sich um eine zufällige Abweichung mit Erwartungswert 0 handelt, wird sich der beobachtete Wert dem tatsächlichen mit einer Erhöhung der Fallzahlen in der Stichprobe immer weiter annähern. Der Sampling Error beschreibt also den nichtsystematischen Einfluss des Prozesses der Stichprobenziehung auf Studienergebnisse. Je weniger Fälle aus der Grundgesamtheit gezogen werden, desto größer ist die Wahrscheinlichkeit, dass die gezogene Stichprobe unsystematisch von der Grundgesamtheit abweicht (Hunter & Schmidt, 2007). Der Sampling Error ist zufallsverteilt um den Erwartungswert Null, das heißt bei einer beliebig großen Stichprobe (die dann die gesamte Grundgesamtheit umfassen würde), wäre er Null, je größer die Stichprobe wird, desto weiter nähert sich der Wert Null an. Es stehen relativ einfache mathematische Verfahren zur Verfügung, um aus einer Meta-Perspektive den Sampling Error von Studien abzuschätzen (vergleiche Hunter & Schmidt, 2007, 81-90). Mathematisch lässt sich die Beziehung folgendermaßen ausdrücken:

$$\delta_e = \frac{(1 - p^2)}{\sqrt{n - 1}}$$

δ_e ist hierbei der Sampling Error, n die Stichprobengröße und p die tatsächliche interessierende Korrelation (abhängige Variable der Metaanalyse) in der Grundgesamtheit. Da p unbekannt ist, ist der Sampling Error so nicht berechenbar, Näherungsverfahren erlauben aber eine Abschätzung.¹⁴² Im Rahmen der in dieser Arbeit vorgelegten Metaanalyse ist der Sampling Error vernachlässigbar. Aus den vorliegenden Daten der Metaanalyse lässt sich als konservative Schätzung ein n von 17.542 (eindeutig unterscheidbare Unternehmensbeobachtungen) ermitteln. Der maximale Sampling Error liegt vor wenn p gegen 0 geht, also nur sehr kleine Korrelationen zwischen Ressourcen und Leistung vorliegen. Für $n = 17.542$ und p gegen 0 liegt δ_e bei rund 0,008 und ist damit vernachlässigbar gering. Auf eine Korrektur für Sampling Error wurde entsprechend im Rahmen der Metaanalyse verzichtet. Die Vergrößerung der Samples, die eine Metaanalyse durch die Zusammenführung von Primärstudien immer durchführt, trägt wesentlich zur Minimierung des Sampling Errors, also der Eliminierung zufälliger Schwankungen um den Erwartungswert von 0 herum, bei (vergleiche auch Crook et al., 2008), ein wesentlicher Vorteil metaanalytischer Verfahren.

2. und 3. Error of measurement in the dependent or independent variable – die abhängige oder die erklärenden Variablen sind schlecht spezifiziert, entweder durch eine allgemein schlechte Erfassung der zu messenden Konstrukte oder durch die unbeabsichtigte Einbeziehung unerklärter Variablen. Dieses Methodenproblem wurde für den Fall der Leistungsoperationalisierung in Abschnitt 5.4.1 ausführlich diskutiert.

4. und 5. Dichotomisierung unabhängiger oder abhängiger Variablen – Ein Problem entsteht, wenn ein komplexes Konstrukt bivariat erfasst werden soll. Hierdurch geht Information verloren, die Erklärungsmacht statistischer Modelle ist darüber hinaus bei bivariaten Operationalisierungen in der Regel schwächer ausgeprägt als bei der Verwendung stetiger Variablen. Dieses Problem ist ein rein methodisches und lediglich durch die Verwendung von Variablen mit mehr Ausprägungen zu lösen. Die Verwendung solcher Ausprägungen ist jedoch nicht immer

¹⁴² Vergleiche auch Hunter & Schmidt, 2007, 81-88 für eine ausführliche mathematische Begründung der heilenden Wirkung großer Stichproben.

möglich. Entsprechend spezialisierte statistische Verfahren wie Logit- und Probitmodelle können hier begrenzt Abhilfe schaffen (vergleiche auch Bowen & Wiersema, 2004). Im Rahmen der in dieser Arbeit vorgelegten Metaanalyse liegen abhängige Variablen mit einer ausreichenden Zahl an Ausprägungen vor, erklärende Variablen im herkömmlichen Sinne gibt es nicht und die Aufteilung des Samples mittels Gruppenvariablen ist in diesem Kontext nicht bedeutsam, da die Gruppierungsvariablen selbst nicht Teil des Erklärungsmodells sind. Das Problem der Dichotomisierung von Variablen ist in der hier vorliegenden Arbeit daher nicht relevant.

6. Reichweitenvariationen der erklärenden Variablen – die statistische Korrelation zwischen Variablen hängt neben dem tatsächlichen Vorliegen eines Zusammenhangs auch davon ab, wie groß die Heterogenität innerhalb einer Variable ausfällt. Ist diese in einer gezogenen Stichprobe im Vergleich zur Grundgesamtheit sehr homogen, sind also die Ausschläge einer Variablen gering, ergeben statistische Modelle keine signifikanten Korrelationen, sind die Ausschläge sehr groß, gehen die Korrelationen gegen 1, ohne dass sich diese Korrelationen auch mit Bezügen in der Grundgesamtheit so einstellen würden. Hunter und Schmidt schlagen zur Eingrenzung dieses Problems den Vergleich der gezogenen Stichprobe mit einer Referenzstichprobe vor. So können Standardabweichungen verglichen werden, fallen diese bei der gezogenen Stichprobe aus dem Rahmen, liegen mathematische Modelle zur Korrektur dieser Abweichungen vor. Da das in der Arbeit gezogene Sample einen wesentlichen Teil der Gesamtheit aller Studien zum Resource-based view abbilden soll, stellt sich dieses Problem im Rahmen der hier durchgeführten Metaanalyse nicht.

7. Reichweitenvariationen der abhängigen Variablen – diese kommen dadurch zustande, dass die abhängige Variable nur einen vorselektierten Teil der Grundgesamtheit misst. Dies ist beispielsweise der Fall, wenn die Leistung von Personalauswahlverfahren untersucht werden soll, jedoch nur Daten für tatsächlich eingestellte Mitarbeiter vorliegen. Besonders gute (abgeworbene oder anderweitig eingestellte) und besonders schlechte (nicht eingestellte oder bereits entlassene) Mitarbeiter sind so nicht Teil des Datensatzes (Beispiel aus Hunter & Schmidt, 2007, 39-41), der so nicht mehr die Gesamtheit der potentiellen Mitarbeiter repräsentiert. Hierbei handelt es sich um die verallgemeinerte Form des Selbstselektionsproblems, auf das weiter unten in Abschnitt A.3.2.2 noch einzugehen sein wird. Hinweise auf Reichweitenprobleme der Variablen im Sample der hier durchgeführten Metaanalyse liegen nicht vor.

8. und 9. Eingeschränkte Konstruktvalidität der abhängigen oder erklärender Variablen – das Problem diskriminante und konvergente Validität zu etablieren ist im Rahmen der Strategischen Managementforschung insbesondere vor dem Hintergrund der Operationalisierung von Leistung relevant, wie im Rahmen dieser Arbeit in Abschnitt 3.1 angedeutet, weitere Ausführungen hierzu finden sich in Steigenberger et al. (2010) sowie Combs et al. (2005). Hunter und Schmidt beschreiben diese Operationalisierungsproblematik insbesondere als imperfekte Korrelation zwischen „intended“ und „study variables“ (Hunter & Schmidt, 2007, 42) im Rahmen der Verwendung von Proxyvariablen für schwer oder kaum direkt messbare Konstrukte. Dieses Problem ist relevant und wurde im Rahmen der Metaanalyse in Abschnitt 5.4.1 durch die Einbeziehung von Korrekturfaktoren berücksichtigt und an verschiedenen Stellen in der Arbeit thematisiert.

10. Technische Fehler – Probleme mit der Eingabe oder Erfassung von Daten. Diese Fehler sind schwer zu spezifizieren und, wie Hunter und Schmidt anmerken, oftmals beachtlich. Vergleiche hierzu auch Starbuck (2004), die die Fehler in den Compustat-Datensätzen analysierten. Aussagen zur Relevanz dieses Problems für die vorgelegte Metaanalyse sind mit den bestehenden Daten nicht möglich.

11. Exogene Faktoren, die erst durch die Studie eingeführt wurden – hierunter fällt beispielsweise der Fall, dass eine Studie Variablen zu unterschiedlichen Zeiten misst und so ein unerwünschter zusätzlicher Faktor das Studienergebnis beeinflusst. Hiergegen hilft lediglich ein entsprechend rigoroses und achtsames Vorgehen bei der Durchführung der Studie. Ein weiteres Phänomen, das in Beziehung mit den Ergebnissen von Studien immer wieder als exogene Einflussgröße genannt wird, ist der Hintergrund des Forschungsprojektes, aus dem eine Studie entstand. Hierunter fallen etwa die Biographie und wissenschaftliche Überzeugung des Forschers, spezifische Interessen des Geldgebers des Forschungsprojektes oder die Art der Publikation (Lipsey, 1994). Im Rahmen dieser Arbeit war es nicht möglich, Aussagen zur Relevanz dieses Effektes zu treffen, regelmäßig werden auch Informationen, die zur Bestimmung dieses Effektes notwendig wären, im Rahmen einer Metaanalyse nicht erfassbar sein.

A.3.2.2 Weitere spezifische Methodenprobleme der Strategischen Managementforschung

Ein methodisches Problem kann im Rahmen von Hypothesentests auftreten, wenn die Trennung zwischen abhängigen und unabhängigen Variablen nicht scharf durchgehalten wird (March & Sutton, 1997; Powell, 2002; Powell, 2003a). Vergleicht man die Leistung von Organisationen ist es notwendig, implizit oder explizit Aussagen über diese Leistung zu treffen, also einen Satz von Kriterien zu entwickeln, der das abstrakte Gebilde der „Leistung“ definieren kann und anhand dessen diese Leistung gemessen werden soll. Ist nun dieser Kriteriensatz nicht vollständig von den eigentlichen Forschungsfragen getrennt, wie etwa im Fall des Resource-based view in der Kritik von Powell (vergleiche hierzu auch die Abschnitte 2.1.3, 2.2.1 und 2.3), wird die Forschungsfrage (im Extremfall) tautologisch, die aus der Theorie abgeleiteten Sätze werden analytisch und die Forschungsarbeit aussagelos. In der Terminologie von Kant: Es ist auf eine saubere Trennung von Subjekt und Prädikat zu achten. Für die Strategische Managementforschung bedeutet dies, dass einer sorgfältigen Definition des Leistungsbegriffs, der zentralen abhängigen Variablen, eine wesentliche Rolle beim Test von Hypothesen und damit Theorien zukommt. Da Erkenntnisfortschritt nur durch derartige Theorietests zu erwarten ist (vergleiche die entsprechenden Ausführungen zur Epistemologie in Abschnitt 2.1), ist dieses Problem prinzipiell im gesamten Forschungsfeld von Relevanz. Unter Umständen könnte eine solche unsaubere Trennung von Prädikat und Subjekt auch die in der Metaanalyse ermittelten hohen Korrelationskoeffizienten zwischen Ressourcen und expertenbasierten Leistungsoperationalisierungen erklären (vergleiche Abschnitt 5.5.2). Über Spekulation hinausgehende Aussagen hierzu sind mit dem hier vorliegenden Forschungsansatz jedoch nicht möglich, weiterer Forschungsbedarf ist hier zu konstatieren.

Hitt et al. (2004) sprechen das Problem der häufig geringen Erklärungsmacht statistischer Modelle an, das auf relativ schlecht spezifizierte Variablen sowie ungeeignete statistische Modellierungen zurückzuführen ist. Ersteres Problem beleuchte ich exemplarisch an der Operationalisierung von Leistung, der wichtigsten Variablen in der Strategischen Management-

forschung (Richard et al., 2009), und bestätige im Wesentlichen die Besorgnis der Autoren. Auf weitere Messprobleme anderer häufiger in der Strategischen Managementforschung genutzter Variablen einzugehen liegt außerhalb des Fokus' dieser Arbeit, die Operationalisierung von Variablen, insbesondere wenn sie komplexe Konstrukte abbilden sollen, stellt in der Sozialforschung sicher eine der zentralen Schwierigkeiten dar (vergleiche auch Abschnitt 2.1.4). Insbesondere die Messung latenter, also nicht direkt beobachtbarer Variablen ist komplex und die Abbildung von Kausalbeziehungen in theoretischer und empirischer Hinsicht schwierig.

Bergh et al. (2004) diskutieren das Problem der „self selection“ (S. 351), das auftritt, wenn die Populationsselektion von der Forschungsfrage nicht unabhängig ist. Dies ist beispielsweise der Fall, wenn besonders gute oder schlechte Unternehmen untersucht werden sollen, oder wenn eine Leistungsoperationalisierung Verwendung findet, die per se bereits einen wesentlichen Teil der Grundgesamtheit der Unternehmen ausgrenzt (etwa bei marktbasieren Kennzahlen, vergleiche Abschnitt 3.2). In diesem Fall besteht zum Einen die Gefahr von Tautologien, wenn der Selektionsalgorithmus zur Bestimmung der Stichprobe von den zu überprüfenden Kausalitäten nicht unabhängig ist (vergleiche oben), zum Zweiten ist es möglich, dass aufgedeckte Kausalitäten sich lediglich aufgrund der Sampleselektion erklären lassen, das Sample damit nicht mehr repräsentativ für die Grundgesamtheit ist. Damit wird zumindest die Aussagekraft des Samples geschwächt (vergleiche auch Abschnitt A.3.2.1, Punkt 7). Das Phänomen der self selection kann im Analyseraster dieser Arbeit als eine besondere Form des Selektionsproblems verstanden werden, das in Abschnitt 3.2 ausführlich diskutiert wurde. Zunächst ist die Begrenzung der Aussagekraft solcher Studien zu beachten, die sich nur auf eine eingeschränkte Teilpopulation beziehen. Der Aufruf zur Rekontextualisierung der Forschungsergebnisse, der als Resümee Kapitel 3.2 beschloss, gilt auch hier. Überdies ist besondere wissenschaftliche Sorgfalt geboten, um ein Tautologieproblem zu vermeiden, das entsteht, wenn Selektionsalgorithmus und erklärende Variablen nicht unabhängig voneinander sind.

Short et al. (2002) weisen auf die Rolle der Stichprobenziehung hin. Die Autoren testen einen Zusammenhang mit vier verschiedenen Sampeln, die mit unterschiedlichen Verfahren aus derselben Grundgesamtheit gezogen wurden. Sie können dabei zeigen, dass sich durch die Wahl der Sampling-Methode eine beliebige Richtung des Zusammenhangs erzeugen lässt. Die Relevanz dieses Methodenproblems kann ich mit dem in dieser Arbeit verfolgten empirischen Ansatz nicht nachprüfen, zur Stichprobentheorie vergleiche beispielsweise Lamnek (1995) oder Hartung et al. (2009).

Das Problem der Auswahl geeigneter statistischer Verfahren und deren teilweise ungenügende Beherrschung durch Autoren, wie von Bergh & Holbein (1997), Bowen & Wiersema (2004), Certo & Semadeni (2006) und anderen diskutiert, ist sicher von Relevanz, allerdings betreffen diese Bedenken das „Handwerkszeug“ des Wissenschaftlers, weniger grundsätzliche Methodenprobleme der Strategischen Managementforschung, so dass eine detaillierte Betrachtung der Stärken und Schwächen verschiedener statistischer Methoden, deren Voraussetzungen und Einsatzfelder der entsprechenden Fachliteratur überlassen werden kann, vergleiche beispielsweise die Buchserie *Research Methodology in Strategy and Management*, die von David J. Ketchen und Donald D. Bergh herausgegeben wird und inzwischen mit fünf Bänden erschienen ist (Ketchen & Bergh, 2004; Ketchen & Bergh, 2005 und folgende).

Danksagung

Für Unterstützung bei der Erstellung dieser Arbeit bin ich folgenden Personen zu Dank verpflichtet: Zuvorderst meinem Doktorvater, Prof. Dr. Peter Pawlowsky, für die Schaffung der Rahmenbedingungen, unter denen diese Arbeit entstehen konnte, und wertvolle Hinweise zu einer ersten Version dieser Arbeit. Frau Nadja Marx für Zuarbeit in ihrer Tätigkeit als studentische Hilfskraft, meinen Kolleginnen Dr. Daniela Menzel und Annett Puggel für Unterstützungen verschiedenster Art und stete Kollegialität, Dr. Peter Mistele für das Fundament, auf dem Teile meiner Überlegungen aufbauen, sowie Prof. Dr. Thorsten Schmidt, Dr. Viktor Slavtchev, Andreas, Tobias und Franz Steigenberger für Kommentare zu früheren Fassungen dieser Arbeit. Und nicht zuletzt Prof. Dr. Michael Fritsch, der in mir das Interesse für wissenschaftliche Fragestellungen geweckt hat.