
3 Prüfungstheorie und Prüfungsforschung

3.1 Prüfungstheorie

Als Realwissenschaft beschäftigt sich die Betriebswirtschaftslehre mit realen Phänomenen der Erfahrungswelt. Zugleich soll die Betriebswirtschaftslehre als anwendungsorientierte Wissenschaft praxisrelevantes Wissen hervorbringen. Insofern besitzt sie gegenüber der Praxis eine dienende Funktion. Diesen Ausführungen folgend ist das Prüfungswesen als Teilbereich der Betriebswirtschaftslehre definitionsgemäß angewandte und Realwissenschaft zugleich.

Demnach verfolgt eine Prüfungstheorie *zwei Ziele*: Das Erkennen und Erklären der Realität sowie deren Gestaltung.

- Das *Erkennen und Erklären der Realität* bezieht sich auf die in der Prüfungsrealität anzutreffenden institutionellen Rahmenbedingungen (→ I.5 und I.6) und die unter deren Einflussnahme herausgebildeten Prüfungspraktiken.
- Die *Gestaltung der Realität* beschäftigt sich vor allem mit der zieladäquaten Ausgestaltung von Normensystemen sowie im Hinblick auf die Prüfungsdurchführung (→ II) damit, wie unter den gegebenen institutionellen Rahmenbedingungen effektiv und effizient vorzugehen ist. Möglich sind auch Vorschläge zur Veränderung der institutionellen Rahmenbedingungen, sofern bestimmte Mittel (Normen) besser geeignet erscheinen, um eine bestimmte Zielsetzung zu erreichen. Ein Beispiel für ein Mittel-Werturteil sind empirische Studien, die belegen, dass der Prüfer unter Verwendung von Fragebögen in signifikant geringerem Maße in der Lage ist, fraud aufzudecken (→ II.4.1.3.2). Hieraus lässt sich die Empfehlung ableiten, in den Prüfungsnormen keine Fragebögen zur Fraud-Aufdeckung aufzunehmen. Das Setzen von Zielen ist indes nicht Gegenstand wissenschaftlicher Untersuchungen, sondern i. S. einer werturteilsfreien Wissenschaft eine politische Entscheidung.

Die Prüfungslehre ist bereits seit Langem als wissenschaftliche Disziplin anerkannt. Allerdings herrscht unverändert keine abschließende Einigkeit hinsichtlich des Begriffs und der Inhalte einer Prüfungstheorie.¹⁴⁸

3.2 Prüfungstheoretische Ansätze

Obwohl keine geschlossene Prüfungstheorie existiert, lassen sich verschiedene *prüfungstheoretische Ansätze* identifizieren. Dabei handelt es sich um gewisse, zumeist nach inhaltlichen und/oder methodologischen Leitideen organisierte natürliche Beurteilungseinheiten.¹⁴⁹ Diese

148 Vgl. hierzu sowie zu den folgenden Ausführungen *Ruhnke* (2000), S. 191 ff. m. w. N.

149 Vgl. z. B. *Schanz* (1988), S. 85 ff., der den Begriff Wissenschaftsprogramm verwendet.

Leitideen werden auch als *Paradigmen* bezeichnet.¹⁵⁰ Paradigmen weisen gegenüber einer Theorie eine geringere Tiefe und Komplexität auf. Forscher, die sich einem bestimmten Paradigma verbunden fühlen, erlangen hierüber Hinweise, in welcher Form sie sich dem Untersuchungsgegenstand nähern können. Demnach geben Paradigmen Anregungen bzw. es werden Regeln aufgestellt, welcher Methoden sich der Forscher bedienen kann oder soll, welche Inhalte konkret untersuchungswürdig sind und aus welchem Blickwinkel diese Inhalte zu betrachten sind. Dabei müssen Paradigmen nicht explizit ausgesprochen werden. Häufig leiten sie die Forschung eher implizit.

Die »Neue Institutionenökonomik«¹⁵¹ ist z. B. ein ursprünglich volkswirtschaftliches Paradigma, welches Anreizwirkungen fokussiert, die von Institutionen ausgehen können. Dabei lassen sich im Rahmen der »Neuen Institutionenökonomik« zwei Hauptströmungen unterscheiden, der Transaktionskostenansatz und der agency-theoretische Ansatz. Diesen beiden theoretischen Ansätzen folgt Abschnitt 2, um die Prüfung von Rechnungslegungsinformationen durch den Wirtschaftsprüfer (→ I.2.1.4.2) und die Existenz des Berufsstands der Wirtschaftsprüfer (→ I.2.3) zu begründen. Auch weitere Teilbereiche der Prüfung sind theoretischen Überlegungen zugänglich. So können theoretische Ansätze auch das Verhältnis des Prüfers zu anderen Akteuren der Prüfungsumwelt erklären. Hier können der agency-theoretische Ansatz ebenso wie der Quasi-Rentenansatz von *DeAngelo* zur Erklärung von Unabhängigkeitsgefährdungen des Prüfers herangezogen werden (→ I.7).

Prüfungstheoretische Ansätze lassen sich danach unterscheiden, ob sie empirisch ausgerichtet sind oder nicht. Empirische oder *erfahrungswissenschaftliche Ansätze* streben das Erkennen und die Erklärung früherer und gegenwärtiger institutioneller Rahmenbedingungen und Prüfungspraktiken sowie eine Prognose des prüferischen Verhaltens an.¹⁵² In hohem Maße erfahrungswissenschaftlich orientiert ist z. B. der auf den Prüfungsprozess bezogene Informationsverarbeitungsansatz (→ II.1.4.2). In Bezug auf die Erklärungsaufgabe geht es darum, das prüferische Verhalten (im Kontext der relevanten Normen und des vorliegenden Prüfungsobjekts) durch die Konstruktion eines geeigneten Systems empirisch gehaltvoller Aussagen zu erklären.¹⁵³ Eine erfahrungswissenschaftliche Theorie fasst wiederum mehrere empirisch bewährte allgemeine Hypothesen (Gesetzesaussagen) systematisch zusammen. Während die Hypothesenherleitung keinen festen Regeln unterliegt, stützt sich die Überprüfung der Hypothesen auf statistische Signifikanztests.

Wie realistisch ist die Herleitung eines solchen Systems empirisch gehaltvoller Aussagen? Kennzeichnend für die Jahresabschlussprüfung sind ihr enger Raum-Zeit-Bezug sowie ihre hohe

150 Vgl. bereits *Kuhn* (1962).

151 Vgl. *Coase* (1937).

152 Vgl. allgemein *Chmielewicz* (1994), S. 150 ff. Zu den einzelnen erfahrungswissenschaftlichen Ansätzen vgl. *Ruhnke* (2000), S. 221 ff.; zum verhaltensorientierten Ansatz vgl. z. B. *Lenz* (2002), Sp. 1924 ff.

153 Zur Umformung von Ursache-/Wirkungszusammenhängen in Ziel-/Mittelbeziehungen im Rahmen einer Prüfungstechnologie vgl. *Ruhnke* (2000), S. 433 ff.

Komplexität. Bereits die regionale Spezifität von Rechnungslegungs- und Prüfungsnormen führt teilweise zu unterschiedlichen Vorgehensweisen bei der Planung und der Durchführung von Prüfungshandlungen. Die anzuwendenden Normen unterliegen ihrerseits kontinuierlichen Veränderungen. Die Herausbildung nomologischer Hypothesen (Gesetzesaussagen, deren Gültigkeit nicht auf ein bestimmtes Raum-Zeit-Gebiet beschränkt ist, sog. Invarianzpostulat¹⁵⁴) zur Erklärung des Prüfungsprozesses scheitert an den angesprochenen Problemen (regionale Spezifität von Normen, die zudem kontinuierlichen Veränderungen unterliegen, und hohe Prüfungskomplexität). Folglich existiert auch keine allgemeingültige Prüfungstheorie, welche in der Lage ist, den Prüfungsprozess auf der Grundlage nomologischer Hypothesen systematisch zu erklären.

Auch bei einer Beschränkung des Invarianzpostulats auf einen bestimmten Raum-Zeit-Ausschnitt (sog. »Quasitheorie«) scheitert die Theoriebildung an der hohen Komplexität. Selbst bei einer Beschränkung auf bestimmte Teilbereiche der Prüfung (z. B. Erklärung des prüferischen Vorgehens bei der Prüfung der Going-concern-Annahme sowie der diesbezüglichen Berichterstattung¹⁵⁵) ist es nicht möglich, ein System abschließender (quasi-)theoretischer Aussagen herzuleiten. Gleichwohl erlauben die hiermit einhergehenden Bemühungen wichtige Einsichten in den Prüfungsprozess. Diese Bemühungen können sich auf die Herausbildung von Bezugsrahmen beziehen, die sich als eine Vorstufe erfahrungswissenschaftlicher Ansätze interpretieren lassen. Ein Bezugsrahmen verfolgt das Ziel, durch gewählte Kategorien das betrachtete Problem verständnisfördernd darzustellen sowie relevante Beziehungen und Mechanismen zwischen diesen Kategorien zu identifizieren. Dabei verfolgt ein Bezugsrahmen zum einen die Aufgabe, Mechanismen für die Integration der reichhaltigen Einzelbefunde bereitzustellen. Zum anderen wird der Anspruch erhoben, künftige Forschungsarbeiten zu steuern.¹⁵⁶

Prüfungstheoretische Überlegungen können auch nicht erfahrungswissenschaftliche Elemente beinhalten (*nicht erfahrungswissenschaftliche Ansätze*). Dies ist z. B. dann der Fall, wenn man versucht, über die Bildung formaler Modelle Erkenntnisse zu gewinnen. Formale Modelle können über realitätsnah gesetzte Annahmen zum Verstehen der Prüfungspraxis beitragen, indem sie das Augenmerk auf zentrale Wirkungsmechanismen lenken. Allerdings sind die gesetzten Annahmen teilweise nur wenig realitätsnah: Beispielsweise unterstellen spieltheoretische Ansätze¹⁵⁷ regelmäßig eine perfekte Prüfungstechnologie, bei deren Einsatz der Prüfer sämtliche Falschdarstellungen beim Mandanten aufdecken kann; dies ist in der Prüfungsrealität gerade nicht der Fall. Als nicht erfahrungswissenschaftliche Ansätze sind neben spieltheoretischen auch formale agency-theoretische Ansätze zu nennen. Auch der messtheoretische Ansatz (→ II.1.4.1) ist regelmäßig nicht erfahrungswissenschaftlich orientiert.

154 Invarianzen sind Regelmäßigkeiten (generelle Aussagen) in der Natur- oder Sozialwelt, deren Geltung nach dem Kriterium der Wahrheit an der Realität zu prüfen ist.

155 Siehe hierzu die Darstellung der Ergebnisse bereits durchgeführter Studien in Carson *et al.* (2013).

156 Vgl. ausführlich Ruhnke (2000), S. 265 ff.; ders. (2003), S. 250 ff.

157 Vgl. z. B. Ewert (2002), Sp. 1908 ff.

3.3 Prüfungsforschung

3.3.1 Einführung

Wie im vorherigen Abschnitt dargelegt, ist eine Prüfungstheorie darauf ausgerichtet, die Prüfungsrealität zu erkennen, zu erklären sowie zu gestalten. *Forschungsbeiträge* umfassen alle systematisch angelegten Beiträge, die einen Erkenntnisfortschritt i. S. der prüfungstheoretisch gesetzten Ziele leisten (→ Abb. I.3-1).

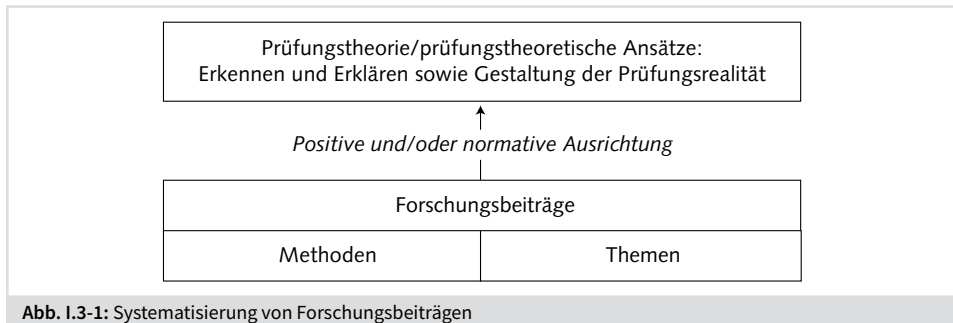


Abb. I.3-1: Systematisierung von Forschungsbeiträgen

Forschungsbeiträge, die sich das Erkennen und Erklären der Prüfungsrealität zum Ziel setzen, haben eine *positive* Ausrichtung. Sie untersuchen etwa das Verhalten von Wirtschaftsprüfern und dieses Verhalten beeinflussende Faktoren. *Normative* Forschung zielt auf die Gestaltung der Realität. Sie gibt Handlungsempfehlungen, sowohl an die Prüfungspraxis, z. B. in Bezug auf die Auslegung von Prüfungsnormen, als auch an Normengeber, in Bezug auf die Gestaltung von Normen, um bestimmte Normierungsziele zu erreichen. Oftmals sind Forschungsarbeiten sowohl positiv als auch normativ ausgerichtet: Beispielsweise bilden in empirischen Studien festgestellte Unzulänglichkeiten bzw. Defizite oftmals Anknüpfungspunkte für Handlungsempfehlungen an den Normengeber.¹⁵⁸

Teilweise initiieren die Normengeber selbst Forschungsarbeiten oder geben Studien heraus, welche die Ergebnisse wissenschaftlicher Forschungsarbeiten zusammenfassen, um den Berufsstand zu informieren oder die Normengebung zu unterstützen.¹⁵⁹ Insbesondere in den

158 Vgl. *Ruhnke/Simons* (2018), S. 73 sowie bereits z. B. *Kupsch* (1985), S. 1164. Zum Einfluss akademischer Arbeiten auf die Prüfungspraxis und die Normengebung siehe auch *AAA Research Impact Task Force* (2009).

159 Vgl. z. B. *IFAC* (2016). Siehe auch <https://www.ifac.org/global-knowledge-gateway/audit-assurance/discussion/accounting-profession-support-academic-research> (Stand: 30.6.2019).

letzten Jahren finden sich Belege für die Verwendung akademischer Arbeiten im Rahmen der Normengebung bzw. Politikberatung (sog. »evidence-based standard setting«).¹⁶⁰

Die positiv und/oder normativ ausgerichteten Forschungsbeiträge lassen sich nach ihrer Methodik sowie nach den behandelten Forschungsthemen unterscheiden. Forschungsmethoden beinhalten systematische Vorgehensweisen, um Erkenntnisse hinsichtlich des Forschungsobjektes (Prüfungswesen) zu erlangen. Das Forschungsobjekt lässt sich wiederum in unterschiedliche Themenfelder systematisieren. Forschungsthemen lassen sich grundsätzlich durch Rückgriff auf unterschiedliche Forschungsmethoden beleuchten. Die Zielsetzung der Forschung (positiv und/oder normativ) und die Wahl der Forschungsmethode sind eng miteinander verknüpft. So setzt das *Erkennen und Erklären der Prüfungsrealität* den Einsatz empirischer Methoden voraus. Dabei bestehen teilweise Restriktionen, da z. B. archivdatenbasierte Untersuchungen nicht anwendbar sind, wenn der Forscher auf vertrauliche Archivdaten nicht zugreifen kann. Hier muss der Forscher ggf. ersatzweise experimentelle Methoden einsetzen.

Empfehlungen zur *Gestaltung der Realität* lassen sich aufgrund empirischer, aber auch logisch-analytischer Überlegungen geben: Beispielsweise kann der Prüfer aufgrund der prinzipiellen Unsicherheit von Prognosen zukunftsorientierter geschätzter Werte (→ II.3.4.3.2.2) eine absolute (100%ige) Prüfungssicherheit nicht gewährleisten. Der Prüfer kann hier lediglich die gesetzten Annahmen auf ihre Plausibilität und die gegenwartsbezogenen Daten auf ihre Richtigkeit prüfen sowie die auf dieser Basis vorgenommene Berechnung des Prognosewerts (geschätzter Wert) nachvollziehen. Insofern handelt der Prüfer auch dann normenkonform, wenn der geschätzte Wert in der Zukunft nicht eintreten sollte. Folglich ist der Prognosefehler kein geeigneter Indikator für die Güte einer Prognose.¹⁶¹ Dies führt zwangsläufig dazu, dass der Prüfer eine 100%ige Prüfungssicherheit auch dann nicht abgeben könnte, wenn die Prüfungsnormen eine solche Sicherheit fordern würden. Die logische Konsistenz eines Normensystems lässt sich demnach nicht über positiv angelegte Forschungsmethoden prüfen.

3.3.2 Forschungsmethoden

Um einen Einblick in den aktuellen Stand der Prüfungsforschung zu geben, werden die Forschungsbeiträge im Folgenden nach Forschungsmethoden und -themen näher untersucht. Bei den angewandten Forschungsmethoden lassen sich empirische und nicht empirische Methoden unterscheiden. Forschungsmethoden beinhalten stets ein systematisches Vorgehen.

¹⁶⁰ Hierzu *Salterio/Hoang/Luo* (2019). Eine Zusammenfassung bestehender, auf die Prüfungsqualität bezogener empirischer Studien in Zusammenhang mit den normativen Änderungsvorschlägen des EU-Grünbuchs bietet *Quick* (2012).

¹⁶¹ Sehr anschaulich hierzu *Tietzel* (1989).

3.3.2.1 Empirische Forschungsmethoden

Empirische Forschungsmethoden haben zum Ziel, Erkenntnisse über die Prüfungsrealität zu erlangen. Hierzu werden *Hypothesen* (research hypotheses) hergeleitet und anhand erhobener Daten überprüft.¹⁶² Die Herleitung von Hypothesen sollte theoriebasiert erfolgen. Insofern stellen Hypothesen ein Bindeglied zwischen Theorie und Empirie dar. Oftmals sind bei der Hypothesenherleitung auch logische Überlegungen, anekdotische Evidenzen oder vorhandene empirische Forschungsergebnisse bedeutsam. Eine Hypothese kann die Richtung des erwarteten Effektes vorgeben (gerichtete Hypothese) oder offenlassen (ungerichtete Hypothese). Fehlt es an ausreichend gesichertem Vorwissen für die Hypothesenbildung, werden Forschungsfragen (research questions) aufgestellt.¹⁶³ Hypothesen sind vor Durchführung des statistischen Tests zu bilden (a priori) und nicht erst, nachdem der Forscher signifikante Korrelationen festgestellt hat (post-hoc).¹⁶⁴

An eine Hypothese ist die Anforderung zu stellen, dass diese empirisch untersuchbar und widerlegbar (falsifizierbar) sein muss. Hypothesen müssen zumindest implizit die Form eines sinnvollen Wenn-dann-Satzes beinhalten. Beispiele für Hypothesen in der Prüfungsforschung sind: »Die Qualität einer externen Prüfung beeinflusst die Qualität der Unternehmensberichterstattung im Anhang positiv« oder »Je größer die Agencykosten eines nicht prüfungspflichtigen Unternehmens sind, desto größer ist die Wahrscheinlichkeit einer freiwilligen Prüfung«.

Werden Hypothesen anhand eines multiplen Regressionsmodells getestet, so wird der Einfluss von unabhängigen Variablen auf eine abhängige Variable untersucht. Die theoretischen Konstrukte (Hypothesen) bedürfen regelmäßig einer empirischen Operationalisierung. Die Unterscheidung zwischen einer konzeptionellen und operativen Ebene ist eng mit dem Begriff der »Libby Boxen« verbunden (→ Abb. I.3-2).¹⁶⁵

162 Vgl. z. B. *Döring/Bortz* (2016), S. 145 ff.; *Eisend/Kuß* (2017), S. 155 ff. m. w. N.

163 Forschungsfragen sind auch dann bedeutsam, wenn sich das Datenmaterial, wie z. B. Interviewdaten, nicht mit statistischen Methoden auswerten lässt.

164 Empirisch lässt sich zeigen, dass große Datenbestände zwangsläufig rein zufällige Korrelationen aufweisen. Hierzu *Calude/Longo* (2016).

165 Vgl. z. B. *Libby/Bloomfield/Nelson* (2002).

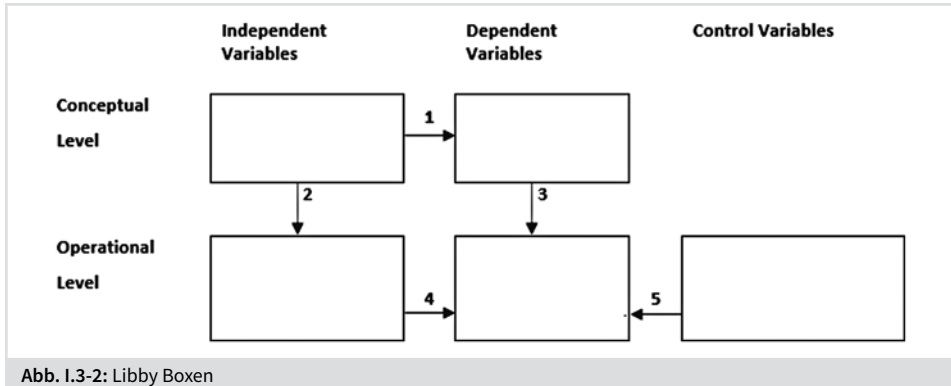


Abb. I.3-2: Libby Boxen

Die Verbindungen zwischen den einzelnen Boxen müssen valide sein. Auf der konzeptionellen Ebene muss der in der Hypothese postulierte Zusammenhang über die Realität eindeutig beschrieben sein (1), so z. B. »Bedrohungen in der Unabhängigkeit des Prüfers beeinflussen die Prüfungsqualität«. Um die Hypothese empirisch zu testen, müssen die abhängige und die unabhängige Variable operational definiert werden. Bedrohungen der Unabhängigkeit des Prüfers werden über Ersatzgrößen wie z. B. die Länge der Mandatsdauer oder einen hohen Anteil an Nicht-Prüfungshonoraren oder Beratungshonoraren gemessen (2). Da die Prüfungsqualität nicht direkt beobachtbar ist, wird diese regelmäßig über Surrogate (sog. Proxy-Variablen)¹⁶⁶ gemessen, so z. B. über den Anteil nicht uneingeschränkt erteilter Bestätigungsvermerke oder den Umfang der Abschlusspolitik beim Mandanten (diskretionäre Periodenabgrenzungen)¹⁶⁷ (3). Dahinter steht die Annahme, dass eine hohe Prüfungsqualität mit einem hohen Anteil nicht uneingeschränkt erteilter Vermerke oder einem niedrigen Umfang diskretionärer Periodenabgrenzungen assoziiert ist.

Der Hypothesentest findet nunmehr auf der operationalen Ebene statt (4). Dabei ist im Hinblick auf den Einfluss von Störgrößen, die erwartungsgemäß einen systematischen Einfluss auf die abhängige Variable haben, zu kontrollieren (5). Als mögliche Kontrollvariablen in dem zuvor skizzierten Beispiel sind die Unternehmensgröße (gemessen z. B. über die Bilanzsumme) und die wirtschaftliche Lage (gemessen z. B. über das Jahresergebnis oder den Cashflow) des

166 Einen Überblick über die Surrogate geben *Maccari-Peukert/Ratzinger-Sakel* (2014). Zur Messung der wahrgenommenen Prüfungsqualität am Kapitalmarkt über implizite Risikoprämien siehe *Qandil* (2014), S. 169 ff.

167 Die Höhe der Periodenabgrenzungen ergibt sich vereinfacht als Differenz zwischen dem Jahresüberschuss und dem Cashflow aus der laufenden Geschäftstätigkeit. Diskretionäre Periodenabgrenzungen sind die Differenz zwischen den so ermittelten Periodenabgrenzungen und den geschätzten normalen Periodenabgrenzungen. Die normalen Periodenabgrenzungen ermitteln sich z. B. vereinfacht anhand eines Durchschnittswertes über mehrere Perioden. Das Jones-Modell berücksichtigt bei der Schätzung der normalen Periodenabgrenzung auch die wirtschaftliche Entwicklung des Unternehmens, z. B. anhand des Umsatzwachstums. Hierzu stellvertr. *Wagenhofer/Ewert* (2015), S. 280 ff.; *Ruhnke/Simons* (2018), S. 358 f. m. w. N.

Mandanten zu nennen. Wichtig ist, dass statistische Signifikanz alleine noch keinen Kausalitätsnachweis darstellt (Korrelation \neq Kausalität). Hierfür bedarf es weiterer Anforderungen, wie z. B., dass der festgestellte Zusammenhang sinnvoll theoretisch fundiert sein muss und dass die Variation der unabhängigen Variablen vor der Variation der abhängigen Variablen eintritt (Grund vor Effekt).¹⁶⁸

Empirische Methoden können in Methoden der Datenerhebung und -auswertung unterschieden werden. Die Methoden der *Datenerhebung* lassen sich wiederum in Archivstudien, Befragungen, Experimente und Fallstudienforschung untergliedern.¹⁶⁹

Archivstudien arbeiten mit Daten, die bereits vorliegen und noch einer formalen Erhebung bedürfen. Oftmals werden Datenbanken, wie z. B. die Dafne-Datenbank oder Audit Analytics, herangezogen. Auch Daten aus dem elektronischen Bundesanzeiger lassen sich hier nutzen. In Betracht kommen ebenfalls andere Archive, z. B. interne Dokumente von Prüfungsgesellschaften, wie etwa Arbeitspapiere des Prüfers oder Listen von festgestellten Prüfungsdifferenzen. Während die externe Validität (Generalisierbarkeit eines empirischen Befundes) solcher Studien regelmäßig hoch ist, kann die interne Validität (kausale Interpretierbarkeit eines empirischen Befundes) bspw. durch eine mangelnde instrumentelle Reliabilität beeinträchtigt sein (z. B. die Messung der Qualität der Abschlussprüfung über mehr oder weniger geeignete Qualitätssurrogate). Als eine besondere Form von Archivstudien sind Metaanalysen interpretierbar, welche mit Hilfe von quantitativen Methoden empirische Befunde (empirische Studien als vorliegende Daten) zu einem bestimmten Problem oder Phänomen zusammenfassen und integrieren (z. B. zur Verwendung von Wesentlichkeitsgrenzen und den Determinanten des audit report lag¹⁷⁰).

BEISPIELE

*Van Linden/Hardies*¹⁷¹ untersuchen den Einfluss der Zugangsvoraussetzungen zum Prüferberuf (\rightarrow I.4.1.1, I.6.5.2.3) auf die Prüfungsqualität. Gezeigt wird, dass ein Masterabschluss (im Vergleich zu einem Bachelorabschluss) sowie eine geforderte praktische Tätigkeit von zumindest drei Jahren die Prüfungsqualität (Abschlusspolitik, gemessen mittels diskretionärer Periodenabgrenzungen) positiv beeinflussen.

Basierend auf Archivdaten, ergänzt durch eine händische Datenerhebung, zeigen *Raffournier/Schatt*¹⁷², dass IFRS-Abschlussersteller (im Vergleich zu Unternehmen, die nach Swiss GAAP bilanzieren) höhere Prüfungshonorare entrichten. Die zugrunde liegende gerichtete Hypothese (\gg Audit fees are higher for firms using IFRS than for those using Swiss GAAPs \ll) wurde unter Hinweis auf die höhere Komplexität der IFRS im Vergleich

168 Ausführlich *Eisend/Kuß* (2017), S. 181 ff. und die dort angegebene Literatur. Zur »fundamental role of theory in drawing causal inferences from empirical evidence« vgl. *Armstrong/Kepler* (2018).

169 Vgl. *Smith* (2017), S. 72 ff.

170 Vgl. *Vance* (2011); *Habib et al.* (2019). Methodisch einführend hierzu *Eisend/Kuß* (2017), S. 211 ff.

171 Vgl. *van Linden/Hardies* (2018).

172 Vgl. *Raffournier/Schatt* (2018).

zu den Swiss GAAP begründet. Kontrolliert wurde u. a. auf die Größe der Abschlussersteller, deren wirtschaftliche Situation, Brancheneffekte sowie auf die Größe der Prüfungsgesellschaft (Big-4-Prüfungsgesellschaft vs. Non-Big-4-Prüfungsgesellschaft). *Ruhnke/Schmitz*¹⁷³ untersuchen (neben den aus der Durchführung von Interviews resultierenden Ergebnissen) auch anhand von internen Dokumenten zum Prüfungsansatz der WPG das Vorgehen bei der Durchführung von präferischen Durchsichten (reviews z. B. gem. ISRS 2400 u. ISRS 2410 bzw. IDW PS 900).

Befragungen können schriftlich, via Post, E-Mail oder Online-Fragebogen erfolgen. Eine besondere Form der Befragung stellen Interviews dar. Hinsichtlich des Grades der Strukturierung kann zwischen einem voll-, halb- und unstrukturierten Vorgehen unterschieden werden. Eine typische Form der Datenerhebung bei einem Interview ist das halbstrukturierte Interview, bei dem der Forscher einem Interview-Leitfaden folgt.¹⁷⁴

Über Befragungen lassen sich z. B. die Erwartungen der Abschlussadressaten an eine Abschlussprüfung oder deren Wahrnehmungen ermitteln, wie sich z. B. eine regulatorische Maßnahme auf die Unabhängigkeit des Abschlussprüfers auswirkt. Auch bei einer Befragung bestehen spezifische Validitätsprobleme. Beispielsweise können Beeinträchtigungen in der Aussagekraft daraus resultieren, dass sich die Antworten der Befragten und ihr tatsächliches Verhalten voneinander unterscheiden oder dass der Befragte das Untersuchungsziel erkennt und hierdurch in seinem Antwortverhalten beeinflusst ist. Weiterhin muss der Forscher berücksichtigen, dass sich über unterschiedliche Darstellungsformen inhaltlich gleicher Fragestellungen das Antwortverhalten beeinflussen lässt (sog. Framing-Effekte, die auch im Hinblick auf die nachstehend beschriebenen Experimente evident sind).

BEISPIELE

Die Erwartungslücke (→ I.3.3.1) ist auf verschiedene Ursachen zurückzuführen, wie z. B. das Versagen der Öffentlichkeit, welches sich wiederum über Erwartungen der Öffentlichkeit messen lässt, die über das Anforderungsniveau der bestehenden Prüfungsnormen hinausgeht. Hier zeigt die Befragung von *Ruhnke/Schmidt*¹⁷⁵ u. a., dass Investoren, Bankenverteter und Aufsichtsratsmitglieder erwarten, dass eine Abschlussprüfung auch eine Prüfung der Ordnungsmäßigkeit der Geschäftsführung (management audit; → I.3.3.6) beinhaltet, obwohl die Prüfungsnormen dies nicht fordern. Eine Erklärung hierfür könnte sein, dass z. B. bei den Aufsichtsräten entgegen der gesetzlichen Vorgabe ein Bedürfnis besteht, sich hinsichtlich ihrer Überwachungstätigkeit durch einen Verweis auf den Abschlussprüfer exkulpieren zu können.

173 Vgl. hierzu *Ruhnke/Schmitz* (2019), insbes. Research Instrument 2.

174 Im Prüfungskontext siehe z. B. *Ruhnke/Schmitz* (2019), insbes. Research Instrument 1.

175 Vgl. *Ruhnke/Schmidt* (2014).

*Weik/Eierle/Ojala*¹⁷⁶ belegen anhand einer Befragung von 405 kleinen Unternehmen, dass der Anteil der nicht in das Management involvierten Unternehmenseigner, die Rechtsform einer AG und die Wichtigkeit von Finanzinformationen die Wahrscheinlichkeit erhöhen, dass ein nicht prüfungspflichtiges Unternehmen sich freiwillig prüfen lässt.

Experimente dienen der Überprüfung von Kausalhypothesen mit dem Ziel, Ursache-Wirkungs-Zusammenhänge zu erforschen.¹⁷⁷ Dabei wird zumindest eine Variable manipuliert und die Reaktion der Teilnehmer untersucht. Experimente können als Labor- oder Feldexperiment ausgestaltet sein, aber auch postalisch oder über das Internet¹⁷⁸ durchgeführt werden. Feldexperimente werden direkt in der Prüfungspraxis, d. h. im Rahmen realer Prüfungsmandate, durchgeführt. Da die zu erfassenden Daten der Prüfer-/Mandatsbeziehung als vertraulich gelten, scheitert diese Form der Datenerhebung in der Regel im Vorfeld. In Laborexperimenten (sowie auch bei einer experimentellen Befragung oder bei einer Durchführung über das Internet) bearbeiten die Teilnehmer hingegen fiktive Prüfungsszenarien. Von Vorteil ist, dass sich Verhaltensgesetzmäßigkeiten der Prüfungspraxis unter streng kontrollierten Bedingungen testen lassen (hohe interne Validität). Oftmals wird getestet, wie der Prüfer in einer bestimmten Entscheidungssituation handelt (abhängige Variable). Über Variationen in der Aufgabenstellung (unabhängige Variablen) wird getestet, ob die vorgenommene Variation das Entscheidungsverhalten beeinflusst. Allerdings lässt sich die Prüfungsrealität naturgemäß nicht perfekt experimentell abbilden. Dies gilt auch im Hinblick auf die Simulation bestimmter Drucksituationen der Prüfungspraxis (z. B. eine mögliche Abhängigkeit des Prüfers vom Mandanten). Insofern stehen in experimentellen Untersuchungen dem Vorteil einer hohen internen Validität stets etwaige Beeinträchtigungen der externen Validität gegenüber.

BEISPIELE

In einem spieltheoretisch motivierten Experiment gehen *Koch/Schmidt*¹⁷⁹ der Frage nach, ob die Verpflichtung zur Offenlegung von Interessenkonflikten, die die Unabhängigkeit des Abschlussprüfers beeinträchtigen, unbeabsichtigte negative Konsequenzen hat. Haben Abschlussprüfer einen Anreiz, abschlusspolitische Maßnahmen zu tolerieren und wird dieser Anreiz offengelegt (unabhängige Variable), erhöht sich in bestimmten Szenarien aufgrund der Offenlegung das Ausmaß der durch den Prüfer tolerierten Abschlusspolitik (abhängige Variable); der Informationsnutzen der Investoren sinkt. Unter Rückgriff auf die sog. »moral licencing theory« wird dies damit erklärt, dass der Prüfer sich durch die Offenlegung des Anreizes für Investoren diesen gegenüber nicht länger in der moralischen Verantwortung sieht. *Ashbar/Ruhnke*¹⁸⁰ zeigen

176 Vgl. *Weik/Eierle/Ojala* (2018).

177 Grundlegend im Prüfungskontext vgl. auch *Trotman* (2001); *Stefani* (2003).

178 Vgl. z. B. *Bryant/Hunton/Stone* (2004).

179 Vgl. ausführlich *Koch/Schmidt* (2010) und die dort angegebene Literatur.

180 Vgl. *Ashbar/Ruhnke* (2019).

experimentell, dass Prüfer die Berichterstattung über Key Audit Matters im Bestätigungsvermerk (→ II.6.3.1) als moralische Lizenz nutzen, um teilweise auf die Buchung von festgestellten Prüfungsdifferenzen (vom Prüfer aufgedeckte Falschdarstellungen, zu deren Korrektur der Mandant aufgefordert wird) zu verzichten.

Fallstudien beschäftigen sich zumeist tiefgehend mit einem oder wenigen Untersuchungsobjekten, dies können z. B. Organisationen, Branchen oder Länder sein.¹⁸¹ Fallstudien in der Prüfungsforschung können etwa den Prüfungsprozess, organisatorische Abläufe in Prüfungsgesellschaften oder auch den Prozess der Entwicklung von Prüfungsstandards einer näheren Betrachtung unterziehen. Da eine Verallgemeinerung der gewonnenen Erkenntnisse aufgrund des individuellen Charakters regelmäßig nicht möglich ist, besitzen Fallstudien häufig explorativen Charakter. Ebenso wie bei Feldexperimenten scheitern Forschungsvorhaben hier häufig im Vorfeld an der Vertraulichkeit der Daten.¹⁸²

BEISPIEL

Die Fallstudie von *Hellman*¹⁸³ untersucht am Beispiel eines börsennotierten Unternehmens in Schweden den Nutzen einer Abschlussprüfung aus Sicht des geprüften Unternehmens. Hierzu wertet er die management letters (→ II.6.3.3) des Abschlussprüfers über mehrere Jahre aus und führt Interviews mit dem Abschlussprüfer und Mitarbeitern des Unternehmens. Über eine fallbezogene Betrachtung einer ausgewählten Prüfer-Mandaten-Beziehung sollen Erkenntnisse dahingehend gewonnen werden, welche Aspekte der Zusammenarbeit mit dem Abschlussprüfer nutzenstiftend für den Mandanten sind und welche Form der Kommunikation zwischen Abschlussprüfer und Unternehmen am geeignetsten ist, um einen hohen Nutzen zu generieren.

Für die Zwecke der *Datenauswertung* lassen sich qualitative und quantitative Methoden einsetzen. Da bestimmte Auswertungsmethoden einen bestimmten Dateninput erfordern, sind Datenerhebung und -auswertung nicht isoliert voneinander zu betrachten. Erhobene Daten werden teilweise qualitativ und quantitativ ausgewertet. Ein Beispiel hierfür ist die Inhaltsanalyse von Dokumenten, die sich z. B. auf Auswertung von comment letters im Normengebungsverfahren oder von Bestätigungsvermerken beziehen kann.¹⁸⁴

Daten aus Archivstudien, schriftlichen Befragungen und Experimenten werden zumeist quantitativ ausgewertet. Im Rahmen von *quantitativen* Auswertungen dient die deskriptive Statistik

181 Vgl. *Smith* (2017), S. 167 ff.

182 Vgl. *Power* (2003), S. 380.

183 Vgl. *Hellman* (2006).

184 Vgl. z. B. zu einer qualitativ und quantitativ ausgerichteten Inhaltsanalyse von comment letters *Yen/Hirst/Hopkins* (2007) sowie die quantitative Inhaltsanalyse von *Smith* (2019), der u. a. die Lesbarkeit von Bestätigungsvermerken mit Hilfe des FOG-Index auswertet. Grundlegend zur Inhaltsanalyse siehe *Döring/Bortz* (2016), S. 540 ff.

der Beschreibung der gewonnenen Daten. Die schließende Statistik (synonym auch inferentielle Statistik) untersucht mittels statistischer Signifikanztests, ob sich die Alternativhypothese stützen lässt. Während die Alternativhypothese das Vorliegen eines Effektes behauptet, verneint die komplementäre Nullhypothese diesen Effekt. Das nachstehende Beispiel verdeutlicht die Bedeutung von Alternativ- und Nullhypothese in Zusammenhang mit statistischen Signifikanztests.

BEISPIEL¹⁸⁵

Im Prüfungskontext könnte die Alternativhypothese behaupten, dass eine Unternehmensgefährdung die Prüfungsdauer verlängert. Die Nullhypothese besagt dann, dass eine Unternehmensgefährdung die Prüfungsdauer nicht verlängert. Da eine abschließende Verifikation der Alternativhypothese nicht möglich ist, setzt die Logik statistischer Signifikanztests an der Nullhypothese an. Dabei führt das Ablehnen der Nullhypothese (signifikante Unterschiede in der Prüfungsdauer bei Unternehmen mit bzw. ohne Unternehmensgefährdung) zu einer Annahme der Alternativhypothese. Das Konstrukt der Unternehmensgefährdung kann z. B. dahingehend operationalisiert werden (siehe hierzu auch die Ausführungen zu den Libby-Boxen weiter oben innerhalb dieses Abschnitts), dass bei einem Cashflow kleiner oder gleich null von einer Gefährdung ausgegangen wird. Das Konstrukt der Prüfungsdauer lässt sich z. B. durch die Anzahl der Tage zwischen dem Abschlusstichtag und dem Datum des Bestätigungsvermerks (sog. audit report lag) operationalisieren.

Wird allein der Effekt der Unternehmensgefährdung auf die Prüfungsdauer untersucht, liegt eine univariate Analyse vor. Um das Wirkungsgeflecht mehrerer einflussnehmender Variablen zu untersuchen, werden multivariate Analysen durchgeführt. Hier lässt sich z. B. mittels einer multiplen Regressionsanalyse zeigen, welche unabhängigen Variablen (Unternehmensgefährdung sowie weitere Variablen wie z. B. Größe des Mandanten und seine Branchenzugehörigkeit) geeignet sind, Variationen in der abhängigen Variable (Prüfungsdauer) zu erklären. Das Bestimmtheitsmaß (erklärte Streuung / Gesamtstreuung) informiert über die Güte der Anpassung der Regressionsfunktion an die empirischen Daten. Beispielsweise gibt ein adjustiertes R^2 von 0,31 an, dass die herangezogenen unabhängigen Variablen 31 % der Variation in der abhängigen Variablen (Prüfungsdauer) erklären.

Fallstudien und qualitative Interviews sind zumeist auf wenige Untersuchungseinheiten begrenzt und aus diesem Grunde oftmals quantitativen Auswertungen nicht zugänglich. Daher werden die erfassten Daten regelmäßig *qualitativ* ausgewertet.¹⁸⁶ Die Daten, welche z. B. in

185 Siehe hierzu *Ruhnke/Au* (1998). Die vorgenommenen Operationalisierungen dienen der Verdeutlichung und werden an dieser Stelle keiner eingehenden kritischen Würdigung unterzogen. Einen Überblick über »audit report lag«-Studien bietet *Abernathy et al.* (2017).

186 Vgl. z. B. *Trompeter/Wright* (2010), die anhand durchgeführter Interviews untersuchen, ob sich der Einsatz analytischer Prüfungshandlungen in der Prüfungspraxis im Zeitablauf geändert hat oder *Ruhnke/Schmitz* (2019), die anhand von explorativ geführten Interviews den Prozess der prüferischen Durchsicht erkunden. Methodisch ähnlich in Bezug auf den Einsatz von Big Data Analytics im Bereich Abschlussprüfung siehe *Salijeni/Samsonova-Taddei/Turley* (2019).

Form von Interviewtranskripten oder Beobachtungsprotokollen vorliegen, können mittels einer qualitativen Inhaltsanalyse ausgewertet werden. Ziel ist es dabei, die Daten in ihrem sozialen Kontext und Bedeutungsumfeld zu interpretieren.¹⁸⁷ Daher geht es in diesen Fällen nicht um den Test von Hypothesen, sondern um Beiträge zur Beantwortung von Forschungsfragen. Diese Methoden sind nicht als minderwertig einzustufen, da bestimmte Forschungsfragen nur qualitativen Methoden zugänglich sind. Zudem bilden mittels qualitativer Methoden erlangte Erkenntnisse oftmals die Voraussetzung, um darauf aufbauend Studien durchzuführen, die quantitative Methoden einsetzen.

3.3.2.2 Nicht empirische Forschungsmethoden

Als nicht empirische Forschungsmethoden lassen sich unterscheiden:

Bei *rein darstellenden* Beiträgen handelt es sich nur dann um Forschungsbeiträge, wenn diese einen Neuigkeitsgehalt aufweisen und nicht schon in derselben oder einer ähnlichen Form in der Literatur zu finden sind. Hierzu zählen auch Beiträge, die einen Sachverhalt aus einem neuen Blickwinkel heraus beschreiben. Auch wenn an die Methodik selbst keine hohen Anforderungen zu stellen sind, lassen sich auf diesem Wege wichtige Erkenntnisbeiträge generieren. Zudem können darstellende Beiträge weitere Forschungsarbeiten initiieren, die sich dann wiederum anderer Methoden bedienen.

Logisch-deduktive Beiträge verwenden deduktive Schlüsse, d. h. sie ziehen Schlussfolgerungen mit Hilfe von Ableitungsregeln aus bestehenden Prämissen.¹⁸⁸ Konkret bedeutet dies, dass z. B. ein bestehender prüfungstheoretischer Ansatz, eine Prüfungsnorm oder auch ein Normierungsziel als Ausgangspunkt genommen wird und hieraus Folgerungen gezogen werden. So kann etwa der Normengeber bei der Frage unterstützt werden, ob sich bestimmte Normierungsziele aufgrund geeigneter Mittel (Prüfungstechniken) in der Prüfungspraxis unter den gegebenen nationalen Umfeldfaktoren umsetzen lassen (sog. realistische Ziel-/Mittelkombinationen).¹⁸⁹ Als logisch-deduktiv sind auch solche Beiträge zu verorten, die sich der juristischen Methodenlehre bedienen und z. B. der Frage nachgehen, ob für die Auslegung eines unbestimmten Rechtsbegriffs ein bestimmter Analogieschluss zulässig ist.

Auch *modelltheoretische* Beiträge bedienen sich logisch-deduktiver Schlüsse. Diese werden hier allerdings in Form mathematisch-formaler Modelle und nicht rein argumentativ gezogen. Modelltheoretische Untersuchungen sind häufig agency- oder spieltheoretisch motiviert und beschäftigen sich etwa mit Entscheidungssituationen und der Wirkung von Anreizen. So wird

¹⁸⁷ Vgl. Döring/Bortz (2016), S. 540 ff.

¹⁸⁸ Vgl. Schanz (1988), S. 44.

¹⁸⁹ Vgl. Mattessich (1995).

z. B. die Wirkung verschiedener Haftungsregimes auf die Sorgfalt des Prüfers bei der Prüfungsdurchführung untersucht.¹⁹⁰

Studiensystematisierungen arbeiten die bestehende Literatur in einem Teilbereich der Prüfungsforschung auf.¹⁹¹ Eng mit der Systematisierung vorhandener Studien ist der *Entwurf eines Bezugsrahmens* verbunden. Ein solcher Entwurf geht über die Systematisierung vorhandener Studien hinaus. Ein Bezugsrahmen erhebt explizit den Anspruch, die vorhandenen Ergebnisse zu integrieren und künftige Forschungsaktivitäten zu steuern, indem z. B. widersprüchliche Forschungsergebnisse oder bisher nicht erforschte Bereiche aufgezeigt werden.¹⁹² Teilweise werden zur Zusammenfassung und Integration empirischer Befunde auch Metaanalysen (→ I.3.3.2.1) durchgeführt, die sich wiederum empirischer Methoden bedienen.

3.3.3 Forschungsaktivitäten und Wissenschaftsbetrieb

Einen Überblick über die Forschungsaktivitäten *deutschsprachiger Prüfungsforscher* in den Jahren 2000 bis 2011 geben *Ruhnke/Schmitz*.¹⁹³ Dabei wurde basierend auf einer Auswertung von 308 Zeitschriftenbeiträgen festgestellt, dass sich oftmals deskriptive Beiträge finden. Allerdings sind die Arbeiten der letzten Jahre zunehmend empirisch ausgerichtet und bedienen sich verstärkt inferenzstatistischer Methoden. Die Forschungsthemen variieren in Abhängigkeit von der Tagesaktualität und werden stark von europäischen und insbesondere nationalen Gegebenheiten geprägt. Am häufigsten finden sich Beiträge zum Themenfeld Prüfungsmarkt und Prüfungsumfeld.

Die Zeitschrift »Die Wirtschaftsprüfung« gibt regelmäßig einen Überblick über die *internationale Prüfungsforschung*. Dabei werden ausgewählte Konferenzbeiträge – z. B. der European Accounting Association (EAA), dem Symposium des European Auditing Research Network (EARNet) und der American Accounting Association (AAA) – synoptisch dargestellt. Von den 180 Beiträgen des ersten Halbjahres 2018 entfielen 66,8% auf Archivstudien und 20,8% auf Experimente sowie 7,1% auf andere empirische Forschungsmethoden. Der Anteil formal-analytischer Beiträge lag bei nur 2,2%.¹⁹⁴ Nicht-empirische Beiträge finden sich kaum.

190 Vgl. *Bigus* (2009), *ders.* (2011).

191 Vgl. z. B. die »research synthesis« von *Carson et al.* (2013) in Bezug auf die Berichterstattung von »Going-concern uncertainties« sowie *Ruhnke/Lee* (2014) in Bezug auf die empirischen Ergebnisse zu den Besprechungen (»brainstorming sessions«) im Prüfungsteam im Rahmen der Aufdeckung von fraud.

192 Zu einem Bezugsrahmen zum Nutzen von Abschlussprüfungen siehe *Ruhnke* (2003) sowie zur »corporate social responsibility assurance« *Velte/Stawinoga* (2017).

193 Vgl. *Ruhnke/Schmitz* (2013).

194 Vgl. *Köhler/Ma* (2018b).

Die im Hinblick auf die Jahre 2012 bis 2018 ausgewerteten Konferenzbeiträge zur internationalen Prüfungsforschung zeigen,¹⁹⁵ dass die Prüfungsforschung unverändert stark Fragestellungen fokussiert, die in einem engen Zusammenhang mit normativen Änderungen stehen. Die Prüfungshonorare stehen oftmals im Mittelpunkt. Im Jahr 2017 fanden sich neben der internen Prüferrotation verstärkt auch Beiträge zur externen Prüferrotation. Erstmals fanden sich auch Forschungsbeiträge zum Bereich (Big) Data Analytics (→ II.5.2.3.3). Im 1. Halbjahr 2018 waren Beiträge zur kritischen Grundhaltung des Abschlussprüfers sowie zur Berichterstattung über Key Audit Matter (KAM) im Bestätigungsvermerk (→ II.6.3.1) verstärkt vertreten.¹⁹⁶ National fokussierte Beiträge sind im Rahmen der Prüfungsforschung unverändert bedeutsam.¹⁹⁷

Ein besonders hohes Potenzial im Hinblick auf künftige Forschungsarbeiten besitzt der Themenbereich Data Analytics betreffend den Einsatzbereich dieser Techniken, etwaige Implikationen für die prüferische Urteilsfindung, die Aus- und Fortbildung, aber auch im Hinblick auf Implikationen in Bezug auf den Prüfermarkt.¹⁹⁸ Unverändert interessante Perspektiven eröffnen u. a. auch die Bereiche Prüfung geschätzter Werte (→ II.3.4.3), Konzernabschlussprüfung (→ II.9.2), Risikobeurteilungen (→ II.3.2.1, II.3.3.1.4), Prüferwesentlichkeit (→ II.1.3), Nutzen einer Prüfung¹⁹⁹ und kritische Grundhaltung des Prüfers (professional skepticism).²⁰⁰

Im Hinblick auf den *Wissenschaftsbetrieb* ist unverändert ein Methodenpluralismus zu konstatieren, auch wenn empirische Arbeiten mittlerweile stark dominieren.²⁰¹ Diese Entwicklung ist vor dem Hintergrund geänderter Rahmenbedingungen des Forschungsbetriebs zu erklären. Zeitschriftenrankings, wie z. B. dem einschlägigen Ranking VHB-JOURQUAL3,²⁰² kommt eine hohe Bedeutung im Rahmen von Berufungsverfahren und Fakultätsevaluationen sowie der Leistungsmittelverteilung innerhalb von Fakultäten zu. Hier hat der Anteil empirischer Beiträge in hochgerankten Journals deutlich zugenommen, was wiederum dazu führt, dass der Prüfungsforscher geneigt ist, sich solcher Forschungsmethoden und -themen zuzuwenden, die eine hochrangige Publikation begünstigen.

In einem engen Zusammenhang hierzu steht die Gefahr der Vernachlässigung von Praxisproblemen, da diese sich teilweise nur mittels Methoden erforschen lassen, die in hochrangigen Journals nicht oder nur sehr eingeschränkt akzeptiert sind, wie z. B. die juristische Methodenlehre

195 Die ausgewerteten Konferenzbeiträge stellen ein erstes Indiz für die thematische Orientierung der internationalen Prüfungsforschung dar. Eine umfassende Bestandsaufnahme steht derzeit aus.

196 Vgl. Köhler/Ma (2018a, 2018b).

197 Vgl. z. B. beschreibend zu den Abschlussprüferhonoraren der DAX30-Unternehmen Quick (2018) und zum Einfluss des Systems der deutschen Corporate Governance auf die Prüferwahl Quick et al. (2018).

198 Hierzu z. B. Köhler/Quick (2018) u. Ruhnke (2019).

199 Zu einem diesbezüglichen Bezugsrahmen siehe Fn. 42.

200 Vgl. McCabe (2018).

201 Zu den folgenden Ausführungen siehe Ruhnke/Schmitz (2013) und die dort angegebene Literatur.

202 VHB-JOURQUAL 3 ist ein Ranking des Verbands der Hochschullehrer für Betriebswirtschaft e. V. (VHB). Hier haben 1.100 Mitglieder des VHB Journals bewertet. Das Ranking ist abrufbar unter <https://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/> (Stand: 30.6.2019). Das derzeitige Ranking basiert auf einer Bewertung aus dem Jahr 2015. Ein erneutes Ranking VHB JOURQUAL 4 ist für 2021 geplant.

oder explorativ angelegte Beiträge. Zudem haben Replikationen, welche der Entdeckung stabiler Wirkungszusammenhänge dienlich sind, regelmäßig keine Aussicht auf eine hochrangige Publikation. Als weiterer Kritikpunkt ist der hochgradig subjektive Begutachtungsprozess für die Aufnahme in ein Journal zu nennen.

Vor dem Hintergrund der Gefahren aus der Verwendung von Rankings ist es zu begrüßen, dass der VHB stets einen verantwortlichen und reflektierten Umgang mit Rankings betont: »VHB-JOURQUAL3 ersetzt also keinesfalls das differenzierte Bild, das sich jeder Kollege und jede Kollegin bei der Beurteilung von Publikationsleistungen machen sollte. Um die Originalität und Relevanz der Fragestellung, die Fundiertheit und argumentative Stringenz des Theoriebeitrags und den methodischen Rigor in Anlage und Ausführung von Modellbildung und/oder Empirie beurteilen zu können, muss man einen Aufsatz tatsächlich lesen – dies war und ist unverändert so.«²⁰³

KONTROLLFRAGEN

1. Wie beurteilen Sie die Möglichkeit, im Prüfungswesen nomologische Hypothesen herzuleiten?
2. Definieren Sie die folgenden Begriffe, grenzen Sie diese voneinander ab und stellen Sie diese in einen sinnvollen Zusammenhang: Bezugsrahmen, erfahrungswissenschaftliche Ansätze, Paradigma, prüfungstheoretische Ansätze und Prüfungstheorie.
3. Gehen Sie kurz auf A-priori- und Post-hoc-Hypothesen ein. Warum gefährden Post-hoc-Hypothesen den Wissenschaftsbetrieb i. S. der Erlangung relevanter Erkenntnisse?
4. Warum sind multiple Regressionsanalysen aussagekräftiger als univariate Analysen?
5. Diskutieren Sie die Möglichkeiten und Grenzen der Erlangung von prüfungstheoretischen Erkenntnissen auf Grundlage der Durchführung von Experimenten.
6. Stellen Sie die empirischen Methoden der Datenerhebung dar. Gehen Sie auch darauf ein, welche Auswertungsmethoden sich für die verschiedenen Arten der Datenerhebung anbieten.

Zitierte Literatur

AAA Research Impact Task Force (2009): The impact of academic accounting research on professional practice: An analysis by the AAA Research Impact Task Force, in: Accounting Horizons, S. 411–456.

203 Hinweise zum verantwortlichen Umgang mit JOURQUAL; hierzu <https://vhbonline.org/vhb4you/jourqual/vhb-jourqual-3/begleitdokumente/hinweise-zum-verantwortlichen-umgang-mit-vhb-jourqual3/> (Stand: 30.6.2019).

- Abernathy, J.L./Barnes, M./Stefaniak, C./Weisbarth, A. (2017): An international perspective on audit report lag: A synthesis of the literature and opportunities for future research, in: *International Journal of Auditing*, S. 100–127.
- Armstrong, C.S./Kepler, J.D. (2018): Theory, research design assumptions, and causal inferences, in: *Journal of Accounting and Economics*, S. 366–373.
- Asbahr, K./Ruhnke, K. (2019): Real effects of reporting key audit matters on auditors' judgment and choice of action, in: *International Journal of Auditing*, S. 165–180.
- Bigus, J. (2009): Abschlussprüferhaftung bei unpräzisen Sorgfaltsstandards und Ambiguitätsaversion, in *Zeitschrift für Betriebswirtschaft*, S. 235–256.
- Bigus, J. (2011): Auditors' liability with overcompensation and reputation losses, in: *OR Spectrum*, S. 287–307.
- Bryant, S.M./Hunton, J.E./Stone, D.N. (2004): Internet-based experiments: Prospects and possibilities for behavioral accounting research, in: *Behavioral Research in Accounting*, S. 107–129.
- Calude, C.S./Longo, G. (2016): *The deluge of spurious correlations in Big Data*, Arbeitspapier, Auckland, Boston.
- Carson, E./Fargher, N./Geiger, M./Lennox, C./Raghunandan, K./Willekens, M. (2013): Audit reporting for Going-concern uncertainty: A research synthesis, in: *Auditing: A Journal of Practice & Theory*, Beiheft 1, S. 353–384.
- Chmielewicz, K. (1994): *Forschungskonzeptionen der Wirtschaftswissenschaft*, 3. Aufl., Stuttgart.
- Coase, R.H. (1937): The Nature of the Firm, in: *Economica*, S. 386–405.
- Döring, N./Bortz, J. (2016): *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*, 5. Aufl., Heidelberg.
- Eisend, M./Kuß, A. (2017): *Grundlagen empirischer Forschung, Zur Methodologie in der Betriebswirtschaftslehre*, Wiesbaden.
- Ewert, E. (2002): Prüfungstheorie, spieltheoretischer Ansatz, in: Ballwieser, W./Coenenberg, A. G. ./Wysocki, K.v. (Hrsg.): *Handwörterbuch der Rechnungslegung und Prüfung*, 3. Aufl., Stuttgart, Sp. 1908–1923.
- Habib, A./Bhuiyan, B.U./Huang, H.J./Miah, M.S. (2019): Determinants of audit report lag: A meta-analysis, in: *International Journal of Auditing*, S. 20–44.
- Hellman, N. (2006): Auditor–client interaction and client usefulness – A Swedish case study, in: *International Journal of Auditing*, S. 99–124.
- IFAC (2016): *The Relationship between Accountancy Expertise and Business Performance, A Review of the Literature*, New York.
- Koch, C./Schmidt, C. (2010): Disclosing conflicts of interest – Do experience and reputation matter?, in: *Accounting, Organizations and Society*, S. 95–107.
- Köhler, A./Ma, W.S. (2018a): Internationale Prüfungsforschung, Überblick über aktuelle Entwicklungen im Jahr 2017, in: *Die Wirtschaftsprüfung*, S. 1003–1008.
- Köhler, A./Ma, W.S. (2018b): Internationale Prüfungsforschung, Überblick über aktuelle Entwicklungen im ersten Halbjahr 2018, in: *Die Wirtschaftsprüfung*, S. 1551–1557.

- Köhler, A./Quick, R. (2018): Editorial: Opportunities for innovative auditing research, in: *International Journal of Auditing*, S. 329–331.
- Kuhn, T.S. (1962): *The structure of scientific revolutions*, Chicago.
- Kupsch, P. (1985): Zum gegenwärtigen Stand des betriebswirtschaftlichen Prüfungswesens, in: *Zeitschrift für Betriebswirtschaft*, S. 1139–1171.
- Lenz, H. (2002): Prüfungstheorie, verhaltensorientierter Ansatz, in: Ballwieser, W./Coenenberg, A. G./Wysocki, K.v. (Hrsg.): *Handwörterbuch der Rechnungslegung und Prüfung*, 3. Aufl., Stuttgart, Sp. 1924–1938.
- Libby, R./Bloomfield, R./Nelson, M.W. (2002): Experimental research in financial accounting, in: *Accounting, Organizations and Society*, S. 775–810.
- Maccari-Peukert, D./Ratzinger-Sakel, N.V.S. (2014): Prüfungsqualität – Eine aktuelle Bestandsaufnahme vor dem Hintergrund internationaler Entwicklungen, in: *Die Wirtschaftsprüfung*, S. 249–257.
- McCabe, S. (2018): CAQ issues request for 2019 audit research, in: *Accounting Today*, November 27.
- Mattessich, R. (1995): Conditional-normative accounting methodology: Incorporating value judgments and means-end relations of an applied science, in: *Accounting, Organizations and Society*, S. 259–284.
- Moser, D.V. (2012): Is accounting research stagnant?, in: *Accounting Horizons*, S. 845–850.
- Power, M.K. (2003): Auditing and the production of legitimacy, in: *Accounting, Organizations and Society*, S. 379–394.
- Qandil, J.S. (2014): Wahrnehmung der Qualität der Abschlussprüfung, Wiesbaden.
- Quick, R. (2012): EC Green Paper proposals and audit quality, in: *Accounting in Europe*, S. 17–38.
- Quick, R. (2018): Abschlussprüferhonorare in den Anhangangaben der DAX30-Unternehmen im Lichte der neuen EU-Verordnung, in: *Betriebs-Berater*, S. 2411–2415.
- Quick, R./Schenk, N./Schmidt, F./Towara, T. (2018): The impact of corporate governance on auditor choice: evidence from Germany, in: *Journal of Management and Governance*, S. 251–283.
- Raffournier, B./Schatt, A. (2018): The impact of International Financial Reporting Standards (IFRS) adoption and IFRS renouncement on audit fees: The case of Switzerland, in: *International Journal of Auditing*, S. 345–359.
- Ruhnke, K. (2000): *Normierung der Abschlußprüfung*, Stuttgart.
- Ruhnke, K. (2003): Nutzen von Abschlussprüfungen: Bezugsrahmen und Einordnung empirischer Studien, in: *Zeitschrift für betriebswirtschaftliche Forschung*, S. 250–280.
- Ruhnke, K. (2019): Auf dem Weg zu Big Data Analytics in der Abschlussprüfung, in: *Die Wirtschaftsprüfung*, S. 64–71.
- Ruhnke, K./Au, K. (1998): Die Dauer der Jahresabschlussprüfung, in: *Die Betriebswirtschaft*, S. 644–658.
- Ruhnke, K./Lee, J.-S. (2014): Besprechungen im Prüfungsteam im Rahmen der Aufdeckung von Fraud im Jahresabschluss, in: *Die Wirtschaftsprüfung*, S. 289–300.
- Ruhnke, K./Schmidt, M. (2014): The audit expectation gap: existence, causes and the impact of changes, in: *Accounting and Business Research*, S. 572–601.
- Ruhnke, K./Schmitz, S. (2013): Prüfungsforschung im deutschsprachigen Raum: Bestandsaufnahme, Entwicklungstendenzen und Herausforderungen, in: *Journal für Betriebswirtschaft*, S. 243–267.

- Ruhnke, K./Schmitz, S. (2019): Review engagements – Structure of audit firm methodology and its situational application in Germany, in: *Journal of International Accounting, Auditing and Taxation*, December, S. 1–17.
- Ruhnke, K./Simons, D. (2018): *Rechnungslegung nach IFRS und HGB*, 4. Aufl., Stuttgart.
- Salijeni, G./Samsonova-Taddei, A./Turley, S. (2019): Big data and changes in audit technology: contemplating a research agenda, in: *Accounting and Business Research*, S. 95–119.
- Salterio, S.E./Hoang, K./Luo, Y. (2019): Communication is a two-way street: Analyzing approaches undertaken to systematically transfer audit research knowledge to policymakers, Arbeitspapier, Kingston, Ontario u. a.
- Schanz, G. (1988): *Methodologie für Betriebswirte*, 2. Aufl., Stuttgart.
- Schmitz, S. (2015): *Der Prozess der prüferischen Durchsicht*, Wiesbaden.
- Smith, M. (2017): *Research methods in accounting*, 4. Aufl., London u. a.
- Smith, K. (2019): Tell me more: A content analysis of expanded auditor reporting In the United Kingdom, Arbeitspapier.
- Stefani, U. (2002): Experimente als Forschungsmethode im Prüfungswesen, in: Richter, M. (Hrsg.), *Entwicklungen der Wirtschaftsprüfung. Prüfungsmethoden – Risiko – Vertrauen*. 4. Symposium der KPMG/Universität Potsdam zur Theorie und Praxis der Wirtschaftsprüfung am 11. und 12.10.2002 in Potsdam, Berlin, S. 243–275.
- Tietzel, M. (1989): Prognoselogik oder: Warum Prognostiker irren dürfen, in: *Jahrbuch für Nationalökonomie und Statistik*, S. 546–562.
- Trompeter, G./Wright, A. (2010): The world has changed – have analytical procedure practices?, in: *Contemporary Accounting Research*, S. 669–700.
- Trotman, K.T. (2001): Design issues in audit JDM experiments, in: *International Journal of Auditing*, S. 181–192.
- Van Linden, C./Hardies, K. (2018): Entrance requirements to the audit profession within the EU and audit quality, in: *International Journal of Auditing*, S. 360–373.
- Vance, D.E. (2011): A meta-analysis of empirical materiality studies, in: *The Journal of Applied Business Research*, S. 53–72.
- Velte, P./Stawinoga, M. (2017): Empirical research on corporate social responsibility assurance (CSRA): a literature review, in: *Journal of Business Economics*, S. 1017–1066.
- Wagenhofer, A./Ewert, R. (2015): *Externe Unternehmensrechnung*, 3. Aufl., Berlin u. a.
- Weik, A./Eierle, B./Ojala, H. (2018): What drives voluntary audit adoption in small German companies?, in: *International Journal of Auditing*, S. 503–521.
- Yen, A.C./Hirst, E./Hopkins, P.E. (2007): A content analysis of the comprehensive income exposure draft comment letters, in: *Research in Accounting Regulation*, S. 53–79.

