

Auf dem Weg zu Kriterien zur Auswahl einer geeigneten Evaluationsmethode für Artefakte der gestaltungsorientierten Wirtschaftsinformatik

Christian Fischer¹

Universität St. Gallen
Institut für Wirtschaftsinformatik
Müller-Friedberg-Str. 8
9000 St. Gallen, Schweiz
christian.fischer@unisg.ch

Abstract: Ziel der gestaltungsorientierten Wirtschaftsinformatik ist die Konstruktion nützlicher Artefakte. Dabei liegt es in der Verantwortung jedes gestaltungsorientiert Forschenden, konstruierte Artefakte zu evaluieren. Dazu werden in der Literatur etliche Forschungsmethoden vorgeschlagen und analysiert. Unsere Literaturanalyse verdeutlicht allerdings, dass es kaum Empfehlungen für Forschende gibt, welche Methode sich für welche Evaluationssituation eignet. In diesem Aufsatz werden daher verschiedene Kriterien diskutiert, die Auswirkungen auf die Auswahl einer Evaluationsmethode haben; weiterhin wird eine begründete Einschätzung abgegeben, unter welchen Voraussetzungen sich welche Evaluationsmethode gut eignet. Es ist davon auszugehen, dass der Beitrag eine nützliche Hilfestellung für die Forschungspraxis darstellt. Eine Schwäche des Beitrags ist allerdings die Subjektivität der Bewertung der Evaluationsmethoden. Zur Verbesserung der intersubjektiven Nachvollziehbarkeit dieser Bewertung ist weitere Forschungsarbeit notwendig.

1 Auswahl einer geeigneten Forschungsmethode als Herausforderung

In der *gestaltungsorientierten* Wirtschaftsinformatik (WI) werden nützliche Lösungen, so genannte Artefakte, für Problemklassen im Gegenstandsbereich der Wirtschaftsinformatik entwickelt [Ba09].² Sie grenzt sich somit von der *erkenntnisorientierten* Wirtschaftsinformatik ab.

¹ Der Autor möchte Herrn Prof. Dr. Ulrich Frank, Universität Duisburg-Essen, und seiner Kollegin Frau Anne Cleven, Universität St. Gallen, für ihre wertvollen Anregungen zum vorliegenden Beitrag recht herzlich danken. Ihre Unterstützung hat maßgeblich die Qualität des Beitrags verbessert.

² Unter dem Begriff Wirtschaftsinformatik werden in dieser Arbeit alle wissenschaftlichen Disziplinen subsummiert, deren Gegenstandsbereich betriebliche Informationssysteme sind, wobei Informationssysteme als soziotechnische Systeme verstanden werden. (Vgl. zum Informationssystembegriff und zum Gegenstandsbereich der Wirtschaftsinformatik z. B. [Le99], [Me09] und [Te99].) Der Begriff Wirtschaftsinformatik umfasst insbesondere alle in diesem Gegenstandsbereich vorherrschenden Forschungsparadigmata und alle Traditionen. (Vgl. zu Forschungstraditionen in der Wirtschaftsinformatik z. B. [Fr08] und [SF07].)

Die erkenntnisorientierte Wirtschaftsinformatik verfolgt im Wesentlichen das Ziel, Phänomene im Gegenstandsbereich der Wirtschaftsinformatik zu beschreiben, zu erklären und/oder vorherzusagen [Gr06]. Die gestaltungsorientierte Wirtschaftsinformatik gleicht hingegen angewandten Wissenschaften wie der Medizin [Ve06a, S. 14 f.], den Ingenieurwissenschaften oder weiten Teilen der Informatik [Gr09]. Eine zentrale Aufgabe der gestaltungsorientierten Wirtschaftsinformatik-Forschung besteht darin, die Nützlichkeit konstruierter Artefakte nachzuweisen (vgl. z. B. [MS95] oder [He04]). March & Smith beschreiben das Ziel einer solchen Artefaktevaluation wie folgt [MS95, S. 258]: „We *evaluate* artifacts to determine if we have made any progress. The basic question is, how well does it work?“ Die Frage nach einer wissenschaftlichen Evaluation eines Artefaktes der gestaltungsorientierten WI ist eng an das mit der Evaluation verbundene Fortschrittsverständnis geknüpft. Unumstritten ist, dass Wissenschaft und Fortschritt untrennbar miteinander verbunden sind [Ku70, S. 162] – auch beruht die normative Legitimation von Wissenschaft, z. B. ihre durch die Gesellschaft anerkannte Förderungswürdigkeit, zu einem großen Teil auf dem vor ihr erwarteten Beitrag zum Fortschritt. Allgemein formuliert, kann als Fortschritt der Übergang von einem Zustand *a* zu einem Zustand *b* definiert werden, wenn *b* besser ist als *a* [Ni09]. Was als *besser* definiert werden und wie bestimmt werden kann, ob ein Forschungsergebnis *besser* ist als ein anderes, ist allerdings Gegenstand der Diskussion – nicht nur in der gestaltungsorientierten Wirtschaftsinformatik, sondern auch in der Wissenschaftstheorie im Allgemeinen [Di07].

Wie Frank herausstellt, ist das Fortschrittsproblem in hohem Maße relevant für die gestaltungsorientierte WI [Fr06]. Die wissenschaftliche Legitimation der Disziplin hängt stark davon ab, (1) wie überzeugend ihr Fortschrittskonzept ist und (2) wie gut sie ihren Beitrag zum Fortschritt ausweisen kann. Ziel dieses Beitrags ist es daher, verschiedene Evaluationsmethoden daraufhin zu untersuchen, (a) unter welchen Bedingungen sie eingesetzt werden können und (b) welche Konsequenzen ihre Anwendung in Bezug auf die Interpretierbarkeit des Evaluationsergebnisses hat.

Dazu ist dieser Beitrag wie folgt gegliedert: Zunächst werden Grundlagen zur gestaltungsorientierten Forschung in der Wirtschaftsinformatik skizziert; verwandte Ansätze werden kurz vorgestellt. In Abschnitt 4 werden dann Empfehlungen für den Einsatz von Evaluationsmethoden, basierend auf Charakteristika verschiedener Evaluationssituationen, herausgearbeitet. Der Aufsatz endet mit einer Zusammenfassung und einer kritischen Reflexion des Beitrags.

2 Grundlagen zur gestaltungsorientierten Wirtschaftsinformatik

2.1 Charakteristika und Ergebnistypen

Vielfach werden drei Charakteristika gestaltungsorientierter Forschung genannt: Nützlichkeit, Innovativität und Allgemeinheit.

Nützlichkeit ist von diesen drei Charakteristika das am häufigsten genannte [Wa92], [MS95], [He04], [Ba09]. Artefakte der gestaltungsorientierten WI sollen relevante Probleme der Praxis lösen; dadurch stiften sie den Akteurinnen und Akteuren, die sie verwenden, Nutzen. Darüber hinaus ist die Innovativität der Forschungsergebnisse ein wichtiges Charakteristikum der Gestaltungsorientierung [MS95], [He04], [Fr06], [Ba09]. Das letzte der obengenannten Charakteristika ist die Allgemeinheit der Forschungsergebnisse [He04], [Fr06], [Ba09]. Forschende haben nicht nur den Anspruch, ein einzelnes, konkretes Problem zu lösen, sondern eine Klasse von Problemen. March & Smith hingegen unterscheiden vier Phasen in einem Forschungsprozess: Bauen, Evaluieren, Theoretisieren (engl.: theorize) und Rechtfertigen, wobei sie nur die ersten zwei Phasen zur gestaltungsorientierten Forschung und die letzten zwei zu den Naturwissenschaften zählen [MS95]. Ihrer Meinung nach wird in der Bauphase eine Lösung für einen konkreten Fall entwickelt; erst beim Theoretisieren werden die Ergebnisse verallgemeinert. Das Attribut „allgemein“ verwenden sie daher ausschließlich in Bezug auf Theorien. Als Ergebnis gestaltungsorientierter Forschung definieren March und Smith vier Artefakttypen: Konstrukte, Modelle, Methoden und Instanzen [MS95].

2.2 Wissenschaftstheoretische Grundlagen

In der gestaltungsorientierten Wirtschaftsinformatik wird das Ziel verfolgt, Wissen zu Problemen, dazugehörigen Problemlösungen und ihrer Nützlichkeit zu gewinnen. Dieses Wissen sollte mit wissenschaftlichen Methoden auf Wahrheit überprüft werden. In der Wissenschaftsphilosophie werden verschiedene Wahrheitsverständnisse diskutiert. Unterschiedliche Wahrheitsverständnissen ziehen unterschiedliche Forschungsmethoden nach sich. Daher werden nachfolgend drei einflussreiche Familien von Wahrheitstheorien dargestellt [Fr06]: (a) die Korrespondenztheorien der Wahrheit, (b) die Kohärenztheorien der Wahrheit und (c) die Konsenztheorien der Wahrheit, vgl. auch Tabelle 1.³

Die Familie der Korrespondenztheorien der Wahrheit definiert Wahrheit als eine Übereinstimmung zwischen einem Wahrheitsträger (engl.: truthbearer), z. B. einer Aussage, und einem Gegenstück, das den Wahrheitsträger wahr macht (engl.: truthmaker), z. B. einem Sachverhalt. (Vgl. als Überblicksarbeiten zur Korrespondenztheorie der Wahrheit z. B. [Da09], [Ki95, S. 119–140].) Frank nennt als Forschungsmethoden, die auf der Korrespondenztheorie der Wahrheit beruhen, das Experiment im Labor oder Feld sowie die Feldstudie [Fr06], darüber hinaus zählt sicherlich auch die positivistische Fallstudie dazu. Die Familie der Kohärenztheorien der Wahrheit nimmt an, dass eine Meinung wahr ist, wenn sie sich kohärent in ein kohärentes System von Meinungen einfügt [Da04, S. 362]. Frank nennt als geeignete Forschungsmethode, die auf der Kohärenztheorie der Wahrheit basiert, die Literaturanalyse.

³ Manche Wahrheitsverständnisse gehen über eine reine Definition des Wahrheitsbegriffs hinaus und geben auch Hinweise darauf, wie eine Aussage als wahr gerechtfertigt werden kann (vgl. zum Verhältnis zwischen Wahrheit und Rechtfertigung z. B. [Ki95, S. 49–54, darüber hinaus auch S. 211–222]). Wir betrachten ausschließlich solche Wahrheitstheorien, die einen Hinweis auf die Verbindung zwischen Wahrheit und Rechtfertigungsmethode geben. Daher schließen wir z. B. Tarskis semantische Wahrheitsdefinition aus, die unter anderem von Becker & Niehaves [BN07] genannt wird.

Aus der Familie der Konsenstheorien der Wahrheit stellen wir exemplarisch den Ansatz dar, den Habermas im Wesentlichen in seinem Aufsatz *Wahrheitstheorien* vorgestellt hat [Ha73] (vgl. für einen Überblick über ähnliche Ansätze [Pu82, S. 142–171]). Dort wird Wahrheit als Ergebnis eines Konsenses definiert, der aus einem idealen Diskurs resultiert. Eine ideale Diskurssituation ist u. a. dadurch gekennzeichnet, (1) dass niemand von dem Diskurs ausgeschlossen wird, der einen Beitrag leisten kann, (2) dass alle Diskursteilnehmer frei sprechen können und insbesondere kein Zwang auf sie ausgeübt wird und (3) dass alle Diskursteilnehmer mit gleicher Stimme sprechen [Ha05, S. 89; zitiert nach BR09]. Letztlich kommt es also darauf an, in einem hierarchiefreien Diskurs auf Basis guter Gründe zu einem Konsens zu gelangen.⁴ Als geeignete Forschungsmethode, basierend auf Habermas Konsenstheorie der Wahrheit, nennt Frank den virtuellen Diskurs, ohne genau zu spezifizieren, was er darunter versteht [Fr06]; sicherlich lassen sich aber auch Fokusgruppen einsetzen, um eine wissenschaftliche Aussage auf Basis der Konsenstheorie zu rechtfertigen (vgl. zum Einsatz von Fokusgruppen in der gestaltungsorientierten WI-Forschung z. B. [Ga07] oder [Tr08]).

Schließlich nennt Frank noch ein viertes Wahrheitsverständnis, die „formale Wahrheit“ (d), der er den formalen Beweis als Methode zuweist [Fr06]. Legt man die Unterscheidung von Immanuel Kant zwischen synthetischen und analytischen Urteilen zugrunde, dann sind die Forschungsmethoden der Korrespondenztheorie, der Kohärenztheorie und der Konsenstheorie der Wahrheit auch zur Bestätigung synthetischer Urteile geeignet, während sich der formale Beweis nur zur Bestätigung analytischer Urteile eignet.

3 Ähnliche Ansätze

In der Literatur zur gestaltungsorientierten Wirtschaftsinformatik wurde eine Vielzahl an Evaluationsmethoden vorgeschlagen. Strukturierungsansätze liefern u. a. Cleven et al. [Cl09], Fettke & Loos [FL03], [FL04], Frank [Fr06], Pfeiffer & Niehaves [PN05], Siau & Rossi [SR07] sowie Wilde & Hess [WH06]. Einige dieser Arbeiten beziehen sich auf die Evaluation von Artefakten, unabhängig von ihrem Typ [Cl09], [Fr06], [PN05], während sich andere ausschließlich auf Artefakte eines bestimmten Typs beziehen, z. B. auf Referenzmodelle [FL03], [FL04] oder Modellierungssprachen [SR07]. Fettke & Loos sowie Siau & Rossi schlagen eine Vielzahl an Evaluationsmethoden vor, beschreiben diese und ordnen sie in einen Bezugsrahmen ein [FL03], [FL04], [SR07]. Pfeiffer & Niehaves schlagen ebenfalls einen Bezugsrahmen zur Evaluation vor, in dem sie den vier Artefakttypen Evaluationskriterien und -methoden zuweisen [PN05]. Frank beschreibt die gestaltungsorientierte Wirtschaftsinformatik in einem Metamodell und definiert darin auch Beziehungen zwischen Elementen gestaltungsorientierter WI-Forschung [Fr06]. Cleven et al. nennen darüber hinaus verschiedene Gestaltungsoptionen bei der Artefaktevaluation in Form eines morphologischen Kastens [Cl09]. Dabei beschreiben sie allerdings weder gegenseitige Abhängigkeiten zwischen verschiedenen Kriterien noch geben sie Empfehlungen, welche Gestaltungsoption unter welchen Bedingungen gewählt werden sollte.

⁴ Es sei hier nur nebenbei bemerkt, dass Habermas in späteren Werken (vgl. z. B. [Ha99]) ein realistisches Wahrheitsverständnis vorgeschlagen hat (vgl. auch [BR09, Abschnitt 3.3]).

Dieser Aufsatz basiert in starkem Maße auf den oben genannten Beiträgen. Allerdings gibt keiner der oben genannten Beiträge umfassende Empfehlungen, in welchen Evaluationssituationen welche Forschungsmethoden geeignet sind. Ein erster Vorschlag für solche Empfehlungen wird in diesem Aufsatz vorgestellt.

4 Gestaltungsempfehlungen zur Auswahl von Evaluationsmethoden

Wir argumentieren, dass die Auswahl einer geeigneten Evaluationsmethode im Wesentlichen von Faktoren aus vier Bereichen abhängt: Zum ersten Bereich gehören diejenigen Restriktionen, die sich aus der Forschungsmethode selbst ergeben, zum zweiten diejenigen vertreten, die sich aus Eigenschaften des Artefakts heraus ergeben, zum dritten diejenigen, die mit dem Evaluationskriterium verbunden sind. Der vierte Bereich beschäftigt sich mit der Interpretierbarkeit der Ergebnisse der Evaluationsmethoden. Es sei darauf hingewiesen, dass die vier Bereiche einen heuristischen Strukturierungsrahmen darstellen, dass sich aber nicht jedes Kriterium trennscharf genau einem Bereich zuordnen lässt. Nachfolgend wird zunächst kurz das forschungsmethodische Vorgehen dieses Beitrags beschrieben (4.1). Im Anschluss daran werden relevante Forschungsmethoden aus der Literatur zusammengetragen (4.2). Schließlich werden die einzelnen Methoden anhand von Kriterien aus den oben genannten Bereichen eingeordnet (4.3 – 4.5 sowie Tabelle 2, Tabelle 3 und Tabelle 4).

4.1 Forschungsmethodisches Vorgehen

Ziel dieses Beitrags ist es, Forschenden in der gestaltungsorientierten Wirtschaftsinformatik eine Hilfestellung bei der Auswahl einer geeigneten Evaluationsmethode zu geben. Konzeptionell orientiert sich dieser Beitrag an den Vorschlägen von Pries-Heje & Baskerville [PB08] (vgl. auch [PB10] sowie ähnliche Ansätze, z. B. [Bu07]), die einen Mechanismus zur Auswahl einer Gestaltungstheorie in Abhängigkeit von spezifischen Rahmenbedingungen und Zielstellungen vorschlagen. Im Zentrum ihres Vorschlags steht eine tabellarische Zuordnung von Merkmalsausprägungen zu geeigneten Gestaltungstheorien. Die Merkmale beziehen sich auf die Einsatzsituation, in diesem Beitrag die Evaluationssituation, und auf die Zielstellung, in diesem Beitrag die Interpretierbarkeit des Evaluationsergebnisses. Eine solche tabellarische Zuordnung findet sich in Tabelle 2, Tabelle 3 und Tabelle 4 dieser Arbeit.

Die Auswahl der Evaluationsmethoden, die in diesem Aufsatz betrachtet werden, ergibt sich aus einer Literaturanalyse. Die Bewertung der einzelnen Forschungsmethoden in Abhängigkeit von verschiedenen Kriterien basiert auf einer umfassenden Sichtung wissenschaftstheoretischer und forschungsmethodischer Literatur. Im Einzelfall ist jede konkrete Bewertung allerdings in hohem Maße subjektiv. Um die Subjektivität dieser Einschätzung zu reduzieren, hat der Autor mit einer Kollegin die Einschätzung diskutiert und infolge der Diskussion an einigen Stellen angepasst.

4.2 Evaluationsmethoden

In Tabelle 1 sind die Evaluationsmethoden, die in den Überblicksarbeiten von Cleven et al. [Cl09], Fettke & Loos [FL03], [FL04], Frank [Fr06] sowie Siau & Rossi [SR07] genannt werden, zusammengefasst und einem Wahrheitsverständnis zugeordnet. Nicht übernommen wurden in der Literatur genannte Evaluationsmethoden, die nicht ein konkretes Vorgehen bei der Evaluation beschreiben, sondern in deren Zentrum die Definition von Evaluationskriterien steht. Dazu zählen bei Fettke & Loos die metrikbasierte, die merkmalsbasierte, die natürlichsprachliche, die paradigmatische sowie die ökonomische Evaluation [FL04], ebenso ein Vorschlag von Frank [Fr07]. Die von Cleven et al. angeführte Forschungsmethode *Prototyp* wurde ebenfalls nicht übernommen, da sie nicht angibt, durch welche Forschungsmethode der Prototyp evaluiert wird [Cl09]. Für die nachfolgende Analyse fassen wir aus der Gruppe (b) in Tabelle 1 die metamodellbasierte, die masterreferenzmodellbasierte und die ontologiebasierte Evaluation zu der Klasse der artefaktvergleichenden Evaluationen zusammen (vgl. zur Anlehnung an Modelle auch [Br07]). Die kognitionspsychologische Evaluation wird als ein Element einer allgemeineren Klasse von Evaluationsmethoden angesehen, nämlich der Klasse theoriebasierten Evaluationsmethoden (vgl. zur Anlehnung an Theorien z. B. [Ge09], [Go04]). In Gruppe (c) wählen wir exemplarisch den virtuellen Diskurs, die Fokusgruppe sowie die Befragung; die Verbalprotokollanalyse sehen wir als eine spezielle Form der Befragung an.

Tabelle 1: Evaluationsmethoden und ihre Wahrheitsverständnisse

Wahrheitsverständnis ([Fr06], s. auch [BN07])	Forschungsmethode
a. Korrespondenztheorie der Wahrheit	Laborexperiment [Cl09], [FL04], [Fr06], [SR07] Feldexperiment [Cl09], [SR07] Feldstudie [FL04], [Fr06] Positivistische Fallstudie [Cl09], [FL04], [SR07] Aktionsforschung [Cl09], [Fr06], [SR07]
b. Kohärenztheorie der Wahrheit	Literaturanalyse [Fr06] Metamodellbasierte Evaluation [FL04] Masterreferenzmodellbasierte Evaluation [FL04] Ontologiebasierte Evaluation [FL04] Kognitionspsychologische Evaluierung [FL04]
c. Konsenstheorie der Wahrheit	Virtueller Diskurs [Fr06] Fokusgruppe (mit Experten) [Ga07], [Tr08] Befragung (von Experten) [Cl09], [FL04], [SR07] Verbalprotokollanalyse [SR07]
d. Analytische Wahrheiten	Formaler Beweis [Cl09], [Fr06]

4.3 Voraussetzungen aus den Evaluationsmethoden

Die einzelnen Evaluationsmethoden stellen unterschiedliche Anforderungen an Evaluationssituationen. Nachfolgend werden Kriterien für solche Anforderungen beschrieben. Gegenstand dieses Abschnitts sind sechs Kriterien, die sich direkt aus der Art der Evaluationsmethoden ergeben: (1) Muss das Artefakt instanziiert werden/sein, damit die Methode angewendet werden kann? (2) Wie groß sollte eine Stichprobe sein, wie viele Fälle sind aufzunehmen? (3) Wie intensiv muss der Kontakt zu Probanden/Fallstudienunternehmen/Interviewpartnern/... etc. sein? Die Bandbreite reicht hier von einem kurzen Ausfüllen eines Fragebogens bis hin zu mehreren persönlichen Kontakten über einen langen Zeitraum hinweg. (4) Setzt die Methode voraus, dass in eine Organisation eingegriffen wird, also dass der Forscher das Unternehmen mitgestalten kann? (5) Setzt die Methode voraus, dass in der Literatur bereits ähnliche Artefakte entwickelt worden sind? (6) Setzt die Methode Theorien voraus, mit denen Gestaltungsentscheidungen bei der Artefaktkonstruktion gerechtfertigt werden können? In Tabelle 2 werden die einzelnen Forschungsmethoden anhand dieser sechs Fragestellungen eingeordnet.

Viele der Forschungsmethoden aus Gruppe (a) haben spezielle Anforderungen an die Instanzierung des Artefaktes (1), die Fallzahl (2), die Kontaktintensität (3) und Eingriffsmöglichkeiten in Unternehmen (4), während sie keinerlei Voraussetzungen an die Verbreitung des Artefaktes in der Literatur stellen (5, 6). Unterschiede zwischen den einzelnen Methoden zeigen sich vor allem bei der Anzahl der benötigten Artefaktinstanzen (1), der Fälle (2) und der Kontaktintensität zur Fallaufnahme (3). Eingriffsmöglichkeiten in Organisationen (4) verlangen nur das Feldexperiment und die Aktionsforschung. Die Feldstudie kann eingesetzt werden, wenn Daten in hoher Stichprobe gesammelt werden können; dafür wird kein intensiver Zugang zu den einzelnen Fällen gefordert. Die Experimente sowie die Fallstudie können zwar auch mit einer geringeren Stichprobenzahl durchgeführt werden, verlangen aber intensiveren Kontakt zu den Probanden/Fallstudienpartnern. Forschungsmethoden aus der Gruppe b, setzen voraus, dass das Artefakt in der Literatur bereits thematisiert wurde (5, 6). Sie sind daher zur Evaluation disruptiver Innovationen, die nicht oder nur wenig auf dem bestehenden Wissenskorpus aufbauen, eher ungeeignet. Methoden, die auf einem Artefaktvergleich basieren, verlangen, dass ähnliche Artefakte bereits publiziert wurden; Methoden, die Gestaltungsentscheidungen durch Theorien rechtfertigen, dass passende Theorien bereits entwickelt wurden. Forschungsmethoden, die auf der Konsenstheorie der Wahrheit beruhen, verlangen nicht, dass ähnliche Arbeiten bereits publiziert wurden. Sie stellen auch geringe Anforderungen an das Vorhandensein von Instanzen. Mit Ausnahme des virtuellen Diskurses verlangen sie zwar nicht viele Interviewpartner/Fokusgruppenmitglieder, aber einen intensiven Kontakt zu diesen Personen. Der formale Beweis stellt keine besonderen Anforderungen an die hier definierten Kriterien.

4.4 Voraussetzungen des zu evaluierenden Artefaktes

Bei der Auswahl einer Evaluationsmethode sollten die Komplexität des Artefaktes (1) und die Art seiner sprachlichen Darstellung berücksichtigt werden (2), vgl. den ersten Teil von Tabelle 3.

Ein formaler Beweis verlangt in der Regel, dass das Artefakt in der Sprache der Mathematik formuliert ist. Die meisten der Methoden des Artefaktvergleichs, die Fettke & Loos nennen [FL04], basieren auf semiformalen Modellen. Eine reine Deduktion von Theorien erfordert eine formalisierte Darstellung der jeweiligen Theorie und des Artefakts. Diejenigen Methoden, die auf der Korrespondenztheorie der Wahrheit basieren, lassen sich generell für Artefakte jeden Formalisierungsgrads verwenden. Artefakte mit hoher Komplexität verlangen, dass die Artefakte den Probanden erklärt werden. Dazu wird ein intensiver Kontakt zu Probanden vorausgesetzt. Die Feldstudie wurde als eine Forschungsmethode charakterisiert, die keinen intensiven Zugang zu den Fällen verlangt. Daher wird sie für komplexe Artefakte eher selten angewandt. Wenn jedoch zu einer großen Zahl an Fällen intensiver Kontakt besteht, kann natürlich auch ein komplexes Artefakt in einer Feldstudie evaluiert werden.

4.5 Voraussetzungen des Evaluationskriteriums

Nicht jede Evaluationsmethode ist für jedes Evaluationskriterium geeignet. Beispielhaft sind die Evaluationskriterien von Aier & Fischer [AF09], [AF10] in Tabelle 3 aufgelistet und bewertet. Zur Evaluation der Nützlichkeit eines Artefaktes eignen sich grundsätzlich alle Methoden mit Ausnahme des formalen Beweises. Allerdings können formale Beweise einen (nicht hinreichenden) Beitrag zu einer Nützlichkeitsevaluation liefern. Die interne Konsistenz eines Artefaktes hingegen kann am besten formal belegt werden. Eher ungeeignet sind die Methoden aus Gruppe (c). Die externe Konsistenz eines Artefaktes, also seine Konsistenz zum Wissenskörper, kann am besten mit Methoden aus Gruppe (b) belegt werden. Unter Umständen kann auch hier auf Methoden aus Gruppe (c) zurückgegriffen werden. Die Anwendungsbreite eines Artefaktes ist im Kern eine Nützlichkeitsevaluation mit hinreichender Anwendungsbreite. Daher ist die Bewertung dieselbe wie bei der Nützlichkeitsevaluation. Die Einfachheit und die Fruchtbarkeit für kommende Forschung lassen sich am ehesten mit einer Methode bestimmen, die auf der Konsens Theorie der Wahrheit basiert.

4.6 Interpretierbarkeit der Ergebnisse

Zum Ausweis von wissenschaftlichem Fortschritt gehört es auch, dass eine Evaluation den Anforderungen der Wissenschaft standhält. Frank nennt fünf Kriterien für Wissenschaftlichkeit [Fr00a]: Wissenschaftliche Aussagen sollen (1) allgemein sein, sich also auf eine Klasse von Sachverhalten beziehen, (2) stabil sein, d. h. zeitlich invariant, (3) intersubjektiv nachprüfbar, (4) bewährt und (5) informativ sein. Davon sind vor allem die Kriterien der Allgemeinheit (1), der Stabilität (2) und der intersubjektiven Nachprüfbarkeit (3) für die Auswahl von Evaluationsmethoden relevant. Die Informativität des Evaluationsergebnisses hängt weniger von der Evaluationsmethode als von dem Gehalt der Hypothese ab, die überprüft werden soll. Die Bewährtheit eines Evaluationsergebnisses hängt nicht von einer einzelnen Evaluation, sondern von mehreren Evaluationsversuchen ab. Die Auswahl einer Evaluationsmethode bei einer Erstevaluation wirkt sich daher nur in geringem Maße auf die Bewährtheit aus.

Tabelle 2: Voraussetzungen der Forschungsmethoden

Voraussetzung der Forschungsmethode	Instanziierung des Artefaktes			Anzahl an Fällen				Kontaktintensität			Eingriff/Untern.		Vergl.-artefakt		Theorie			
	keine	eine/wenige	vielen	keine	eine(r)	wenige	vielen	kein Zugang	Z. nicht intensiv	Zugang intensiv	nein	ja	nicht vorhanden	vorhanden	nicht vorhanden	vorhanden	nicht vorhanden	
Forschungsmethode	Laborexperiment	●	●	●	●	●	●	○	●	●	●	●	●	●	●	●	●	●
	Feldexperiment	○	●	●	●	●	●	○	○	●	○	○	○	○	○	○	○	○
	Feldstudie	○	○	●	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	Positiv. Fallstudie	○	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○
	Aktionsforschung	●	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○
(a)	Artefaktvergleich	●	●	●	●	●	●	○	○	○	○	○	○	○	○	○	○	○
	Theoriebas. Begründung	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
(b)	Virtueller Diskurs	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	Expertenbefragung	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
(c)	Fokusgruppe	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	Formaler Beweis	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
(d)	Formaler Beweis	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○
	Formaler Beweis	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Legende: Die Methode ist bei dieser Ausprägung ○ ungeeignet, nur ○ in wenigen Ausnahmefällen geeignet, ● mit großen Einschränkungen geeignet, ● mit kleinen Einschränkungen geeignet, gut geeignet ●. Das Merkmal ist für diese Methode irrelevant ⊗.

Tabelle 3: Voraussetzungen des Artefaktes und des Kriteriums

Voraussetzungen des Artefaktes/ des Kriteriums		Artefakt				Kriterium						
		Komplexität		Formalisierungsgrad		Nützlichkeit	Konsistenz	interne Konsistenz	extreme breite	Anwendungs-	Einfachheit	Fruchtbark. f. w. Forschung
Forschungsmethode		gering	hoch	naturlich-sprachlich	semi-formal							
(a)	Laborexperiment	●	●				●	○	○	●	◐	◐
	Feldexperiment	●	●				●	○	○	●	◐	◐
	Feldstudie	●	◐	⊗	⊗	⊗	●	○	○	●	◐	◐
	Positiv. Fallstudie	●	●				●	○	○	●	◐	◐
	Aktionsforschung	●	●				●	○	○	●	◐	◐
(b)	Artefaktvergleich	●	●	◐	●	◐	●	◐	●	●	◐	◐
	Theoriebas. Begründung	●	●	◐	◐	●	●	○	●	●	◐	◐
(c)	Virtueller Diskurs	●	●	●	●	●	●	◐	◐	●	◐	◐
	Expertenbefragung	●	●	●	◐	◐	●	◐	●	●	●	●
	Fokusgruppe	●	●	●	◐	◐	●	◐	●	●	●	●
(d)	Formaler Beweis	●	●	○	◐	●	◐	●	◐	◐	○	○

Legende: s. Tabelle 2

In Tabelle 4 werden die einzelnen Methoden in Bezug auf die drei Kriterien bewertet. Die intersubjektive Überprüfbarkeit hängt im Wesentlichen von der Dokumentation ab. Am besten lassen sich literaturbasierte und formale Evaluationen nachvollziehen. Die zu erwartende Stabilität einer Evaluation hängt bei den quantitativ-empirischen Methoden von der Fallzahl und den statistischen Gütekriterien ab, bei den qualitativ-empirischen Methoden ebenfalls von der Fallzahl und der Qualität der Analyse. Bei den literaturbasierten Verfahren hängt die zu erwartende Stabilität eines Evaluationsergebnisses einerseits von der Stabilität des Bezugswissens aus der Literatur ab, andererseits davon, wie stark sich an Bezugswissen angelehnt wird. Analogieschlüsse sind z. B. schwache Formen der Anlehnung, eine formale Deduktion ist hingegen eine starke Form. Die zu erwartende Stabilität eines Evaluationsergebnisses, das auf der Konsenstheorie der Wahrheit basiert, hängt stark von der Erfahrung der Experten und von der Qualität der Rahmenbedingungen des Diskurses ab. Es ist davon auszugehen, dass die Expertenmeinung präziser ist, wenn das zu evaluierende Artefakt durch einen Prototypen instanziiert ist. Letztlich ist es wichtig, dass alle Diskursteilnehmer dasselbe Problemverständnis haben.

Die Generalisierbarkeit der Evaluationsergebnisse hängt bei den empirischen Methoden von der Breite der Stichprobe ab, also nicht allein von der Stichprobenzahl. Die Generalisierbarkeit eines literaturbasierten Evaluationsergebnisses hängt auch von der Anwendungsbreite des referenzierten Teils des Wissenskorpus ab. Ggf. ist es auch möglich, sich an verschiedenen, ähnlichen Quellen anzulehnen. Bei Evaluationen auf Basis der Konsenstheorie der Wahrheit hängt die Generalisierbarkeit von der Erfahrung der Experten ab. Vor allem, indem Experten aus unterschiedlichen Bereichen hinzugezogen werden, lässt sich die Generalisierbarkeit des Evaluationsergebnisses steigern.

6 Zusammenfassung und Diskussion

Dieser Aufsatz gibt Empfehlungen zur Auswahl einer Forschungsmethode zur Evaluation eines Artefaktes der gestaltungsorientierten Wirtschaftsinformatik. Obwohl dieser Beitrag auf eine Vielzahl ähnlicher Arbeiten aufbauen kann, wurden bisher noch nicht derart umfassende Empfehlungen gegeben. Es ist damit zu rechnen, dass dieser Beitrag einen hohen Nutzen für die Forschungspraxis der gestaltungsorientierten WI hat. Die Grenzen des Beitrags liegen in der Subjektivität der Bewertung der einzelnen Forschungsmethoden. Dies gilt insbesondere für die Bewertung der Interpretierbarkeit von Evaluationsergebnissen. Um die Qualität der Bewertung zu steigern, könnte bspw. eine Fokusgruppe mit erfahrenen gestaltungsorientiert Forschenden aus der Wirtschaftsinformatik gebildet werden, in deren Rahmen die einzelnen Kriterien und ihre Ausprägungen diskutiert werden. Eine solche Evaluation ist Gegenstand künftiger Forschung.

Literatur

- [AF09] Aier, S.; Fischer, C.: Scientific Progress of Design Research Artefacts. In: Proceedings of the 17th European Conference On Information Systems 2009.
- [AF10] Aier, S.; Fischer, C.: Criteria of Progress for Information Systems Design Theories. In: Information Systems and E-Business Management, Online First: DOI:10.1007/s10257-010-0130-8.
- [Ba09] Baskerville, R. L.; Pries-Heje, J.; Venable, J.: Soft design science methodology. In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2009), New York 2009, S. 1–11.
- [BN07] Becker, J.; Niehaves, B.: Epistemological perspectives on IS research: a framework for analyzing and systematizing epistemological assumptions. In: Information Systems Journal 17 (2007), S. 197–214.
- [Br07] vom Brocke, J.: Design Principles for Reference Modeling – Reusing Information Models by Means of Aggregation, Specialization, Instantiation, and Analogy. In: Fettke, P.; Loos, P. (Hrsg.): Reference Modeling for Business Systems Analysis. Idea Group, Hershey 2007, S. 47–75.

Tabelle 4: Interpretierbarkeit des Ergebnisses

Interpretationsziel Forschungsmethode	Intersubjektive Überprüfbarkeit	Stabilität	Allgemeinheit
(a)	Laborexperiment	● - ●	<ul style="list-style-type: none"> ● Breite der Stichprobe
	Feldexperiment	●	
	Feldstudie	●	
	Positiv. Fallstudie	● - ●	
	Aktionsforschung	●	
(b)	Artefaktvergleich	● - ●	<ul style="list-style-type: none"> ● Anwendungsbreite der Aussagen aus der Literatur
	Theoriebas. Begründung	<ul style="list-style-type: none"> ● Stabilität des Bezugswissens ● Stärke der Anlehnung an Bezugswissen (Deduktion – Analogie) 	
(c)	Virtueller Diskurs	●	<ul style="list-style-type: none"> ● Erfahrung der Experten
	Expertenbefragung	●	
	Fokusgruppe	●	
(d)	Formaler Beweis	●	●

Legende: s. Tabelle 2.

- [BR09] Bohman, J.; Rehg, W.: Jürgen Habermas. In: Zalta, E. N. (Hrsg.): The Stanford Encyclopedia of Philosophy 2009.
- [Bu07] Bucher, T. et al.: Situational Method Engineering – On the Differentiation of "Context" and "Project Type". In: Proceedings of the IFIP WG8.1 Working Conference on Situational Method Engineering – Fundamentals and Experiences (ME07), Geneva 2007, S. 33–48.
- [BW08] Bucher, T.; Winter, R.: Dissemination and Importance of the „Method“ Artifact in the Context of Design Research for Information Systems. In: Proceedings of the Third International Conference on Design Science Research in Information Systems and Technology (DESRIST 2008), Atlanta, GA 2008, S. 39–59.
- [CI09] Cleven, A.; Gubler, P.; Hüner, K.: Design Alternatives for the Evaluation of Design Science Research Artifacts. In: Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology 2009.
- [Da04] David, M.: Theories of Truth. In: Niiniluoto, I.; Sintonen, M.; Woleński, J. (Hrsg.): Handbook of Epistemology. Kluwer, Dordrecht 2004, S. 331–414.
- [Da09] David, M.: The Correspondence Theory of Truth. In: Zalta, E. N. (Hrsg.): The Stanford Encyclopedia of Philosophy 2009.
- [Di07] Dilworth, C.: Scientific Progress. 4 Aufl., Springer, Dordrecht, NL 2007.
- [FL03] Fettke, P.; Loos, P.: Multiperspective Evaluation of Reference Models – Towards a Framework. In: Jeusfeld, M. A.; Pastor, Ó. (Hrsg.): Proceedings of the ER 2003 Workshop, Springer, Berlin 2003, S. 80–91.
- [FL04] Fettke, P.; Loos, P.: Entwicklung eines Bezugsrahmens zur Evaluierung von Referenzmodellen. Arbeitsbericht der Forschungsgruppe Information Systems and Management der Johannes Gutenberg-Universität Mainz, Mainz 2004.
- [Fr00a] Frank, U.: Evaluation von Artefakten in der Wirtschaftsinformatik. In: Heinrich, L. J.; Häntschel, I. (Hrsg.): Evaluation und Evaluationsforschung in der Wirtschaftsinformatik – Handbuch für Praxis, Lehre und Forschung. Oldenbourg, München 2000, S. 35–48.
- [Fr06] Frank, U.: Towards a Pluralistic Conception of Research Methods in Information Systems Research. Forschungsbericht des Instituts für Informatik und Wirtschaftsinformatik der Universität Duisburg-Essen, Essen 2006.
- [Fr07] Frank, U.: Evaluation of Reference Models. In: Fettke, P.; Loos, P. (Hrsg.): Reference Modeling for Business Systems Analysis. Idea Group 2007, S. 118–140.
- [Fr08] Frank, U.; Schauer, C.; Wigand, R. T.: Different Paths of Development of Two Information Systems Communities: A Comparative Study Based on Peer Interviews. In: Communications of the AIS 22 (2008) 21, S. 391–412.
- [Ga07] Gibson, M.; Arnott, D.: The Use of Focus Groups in Design Science Research. In: Proceedings of the 18th Australasian Conference on Information Systems, Toowoomba, Australia, 2007.
- [Ge08] Gericke, A.: Konstruktionsforschung und Artefaktkonstruktion in der gestaltungsorientierten Wirtschaftsinformatik: Ein Literaturüberblick. Forschungsbericht des Instituts für Wirtschaftsinformatik BE HSG/IWI1/1, Universität St. Gallen, St. Gallen 2008.

- [Ge09] Gehlert, A. et al.: Towards a research method for theory driven design research. In: *Business Services: Konzepte, Technologien, Anwendungen*; Proceedings der 9. Internationalen Tagung Wirtschaftsinformatik, Wien 2009, S. 441–450.
- [GJ07] Gregor, S.; Jones, D.: The Anatomy of a Design Theory. In: *Journal of the Association for Information Systems* 8 (2007) 5, S. 312–335.
- [Go04] Goldkuhl, G.: Design Theories in Information Systems – A Need for Multi-Grounding. In: *JITTA: Journal of Information Technology Theory and Application* 6 (2004) 2, S. 59–72.
- [Gr06] Gregor, S.: The Nature of Theory in Information Systems. In: *MIS Quarterly* 30 (2006) 3, S. 611–642.
- [Gr09] Gregor, S.: Building Theory in the Sciences of the Artificial. In: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*, New York 2009.
- [GW09] Gericke, A.; Winter, R.: Entwicklung eines Bezugsrahmens für Konstruktionsforschung und Artefaktkonstruktion in der gestaltungsorientierten Wirtschaftsinformatik. In: *Becker, J.; Krcmar, H.; Niehaves, B. (Hrsg.): Wissenschaftstheorie und gestaltungsorientierte Wirtschaftsinformatik*. Physica, Heidelberg 2009, S. 195–210.
- [Ha05] Habermas, J.: *Zwischen Naturalismus und Religion*. Suhrkamp, Frankfurt/Main 2005.
- [Ha73] Habermas, J.: *Wahrheitstheorien*. In: *Fahrenbach, H. (Hrsg.): Wahrheit und Reflexion*. Günter Neske Verlag, Pfullingen 1973, S. 211–265.
- [Ha99] Habermas, J.: *Wahrheit und Rechtfertigung*. Suhrkamp, Frankfurt/Main 1999.
- [He04] Hevner, A. R. et al.: Design Science in Information Systems Research. In: *MIS Quarterly* 28 (2004) 1, S. 75–105.
- [Ki95] Kirkham, R. L.: *Theories of Truth*. MIT Press, Cambridge, Massachusetts 1995.
- [Ku70] Kuhn, T. S.: *The Structure of Scientific Revolutions*. 3. Aufl., Chicago University Press, Chicago 1970.
- [Le99] Lee, A. S.: Inaugural Editor's Comments. In: *MIS Quarterly* 23 (1999) 1, S. v–xi.
- [Me09] Mertens, P.: Was ist Wirtschaftsinformatik? In: *Mertens, P. et al. (Hrsg.): Studienführer Wirtschaftsinformatik 2009/2010: Studieninhalte, Anwendungsfelder, Berufsbilder*. Gabler, Wiesbaden 2009, S. 1–8.
- [MS95] March, S. T.; Smith, G. F.: Design and Natural Science Research on Information Technology. In: *Decision Support Systems* 15 (1995) 4, S. 251–266.
- [Ni09] Niiniluoto, I.: Scientific Progress. In: *Zalta, E. N. (Hrsg.): The Stanford Encyclopedia of Philosophy* 2007.
- [Of10] Offermann, P. et al.: Artifact Types in Information Systems Design Science – A Literature Review. In: *Proceedings, Global Perspectives on Design Science Research*, Berlin 2010.
- [PB08] Pries–Heje, J.; Baskerville, R. L.: The Design Theory Nexus. In: *MIS Quarterly* 32 (2008) 4, S. 731–755.
- [PB10] Pries–Heje, J.; Baskerville, R. L.: Management Design Theories. In: *Proceedings, Human Benefit through the Diffusion of Information Systems Design Science Research*, Boston 2010, S. 263–281.

- [PN05] Pfeiffer, D.; Niehaves, B.: Evaluation of Conceptual Models: A Structuralist Approach. In: Proceedings, Proceedings of the 13th European Conference on Information Systems (ECIS 2005), Regensburg 2005.
- [Pu82] Puntel, L. B.: Wahrheitstheorien in der neueren Philosophie. 2 Aufl., Wissenschaftliche Buchgesellschaft, Darmstadt 1982.
- [SF07] Schauer, C.; Frank, U.: Wirtschaftsinformatik und Information Systems – Ein Vergleich aus wissenschaftstheoretischer Sicht. In: Lehner, F.; Zelewski, S. (Hrsg.): Wissenschaftstheoretische Fundierung und wissenschaftliche Orientierung der Wirtschaftsinformatik. Gito, Berlin 2007, S. 121–154.
- [SR07] Siau, K.; Rossi, M.: Evaluation Techniques for Systems Analysis and Design Modelling Methods – A Review and Comparative Analysis. In: Information Systems Journal 49 (2007) 5, S. 455–474.
- [Su09] Sunyaev, A.; Hansen, M.; Krcmar, H.: Method Engineering: A Formal Description. In: Papadopoulus, G. A. et al. (Hrsg.): Information Systems Development. Towards a Service Provision Society. Springer, New York, Dordrecht, Heidelberg, London 2009, S. 645–654.
- [Te99] Teubner, A.: Organisations- und Informationssystemgestaltung: Theoretische Grundlagen und integrierte Methoden. Gabler, Wiesbaden 1999.
- [Th06] Thomas, O.: Das Referenzmodellverständnis in der Wirtschaftsinformatik: Historie, Literaturanalyse und Begriffsexplikation. Veröffentlichungen des Instituts für Wirtschaftsinformatik im Deutschen Forschungszentrum für Künstliche Intelligenz 187, Institut für Wirtschaftsinformatik im Deutschen Forschungszentrum für Künstliche Intelligenz, Saarbrücken 2006.
- [Tr08] Tremblay, M. C.; Hevner, A. R.; Berndt, D. J.: The Use of Focus Groups in Design Science Research. In: Proceedings, Proceedings of the Third International Conference on Design Science Research in Information Systems and Technology (DESRIST 2008), Atlanta, GA 2008, S. 17–37.
- [Va06] Vahidov, R.: Design Researcher's IS Artifact – A Representational Framework. In: Proceedings, Proceedings of the First International Conference on Design Science Research in Information Systems and Technology (DESRIST 2006) 2006, S. 19–33.
- [Ve06a] Venable, J.: The Role of Theory and Theorising in Design Science Research. In: Proceedings, Proceedings of the 1st International Conference on Design Science in Information Systems and Technology (DESRIST 2006), Claremont, CA 2006, S. 1–18.
- [Wa92] Walls, J. G.; Widmeyer, G. R.; El Sawy, O. A.: Building an Information System Design Theory for Vigilant EIS. In: Information Systems Research 3 (1992) 1, S. 36–59.
- [WH06] Wilde, T.; Hess, T.: Methodenspektrum der Wirtschaftsinformatik: Überblick und Portfoliobildung. Arbeitsbericht 2/2006 des Instituts für Wirtschaftsinformatik und Neue Medien der Ludwig-Maximilians-Universität München. München 2006.
- [Wi09] Winter, R.; Gericke, A.; Bucher, T.: Method versus Model – Two Sides of the Same Coin? In: Proceedings, Advances in Enterprise Engineering III: 5th International Workshop, CIAO! 2009, and 5th International Workshop, EOMAS 2009, held at CAiSE 2009, Amsterdam, The Netherlands, June 8–9, 2009, Springer, Berlin 2009, S. 1–15.

