



# ESSENER KONTRA.

## Wissenschaft und Praxis im Dialog

### **Kill Thrill-Index – Warum das Ziel weit verfehlt wird**

Kritische Anmerkungen zu der Studie „Category Killer“ von der Unternehmensberatung Dr. Wieselhuber & Partner

von Hendrik Schröder, Universität Duisburg-Essen, Campus Essen

Im Frühjahr 2014 hat die Unternehmensberatung *Dr. Wieselhuber & Partner* eine Studie mit dem Titel „Category Killer – Der stationäre Handel unter Zugzwang“ herausgegeben. Die empirische Untersuchung wurde in Zusammenarbeit mit *Facit Research* durchgeführt, als Autoren werden *Timo Renz*, *Johannes B. Berentzen* und *Jan-F. Stahlbock* genannt.

Die Studie will untersuchen, welche Einzelhändler am besten auf den Online-Handel eingestellt sind und welche davon besonders bedroht sind. Für den ersten Bereich wird ein „Fit Score-Index“ erstellt, für den zweiten ein „Kill Thrill-Index“. Wir befassen uns hier mit dem Kill Thrill-Index, der anzeigen soll, „wie gefährdet die stationären Anbieter durch potenzielle Category Killer aus der Online-Welt sind“ (Renz/Berentzen/Stahlbock 2014, S. 4). Unter Category Killer verstehen die Autoren solche Händler, die mit ihren Online-Shops stationäre Händler aus dem Markt gedrängt haben bzw. drängen: „Amazon hat es vorgemacht und als Erstes eine ganze Kategorie, gekillt!“ (S. 3; soweit nicht anders angegeben, beziehen sich alle Seitenangaben auf Renz/Berentzen/Stahlbock 2014)

In ihren Schlussfolgerungen kommen die Autoren zu folgenden Aussagen: „Der Kill Thrill gibt erste Hinweise auf eine latente Bedrohung durch einen Category Killer.“ (S. 30) „Eine hohe Bedrohung muss zwar nicht zwingend einer gegenwärtig erfolgreichen Position widersprechen. Kommt jedoch mangelnde ‚Fitness‘ der Händler einer Branche hinzu, sind die Voraussetzungen für potenzielle Category Killer günstig.“ (S. 30) Namentlich werden 50 Händler aufgeführt, für die die Unternehmensberatung den so genannten Kill Thrill ermittelt hat. Die maximale Bedrohung liegt bei 100 Punkten. Den ersten zehn Plätzen werden zugewiesen: *Vedes* (84,0), *Medimax* (81,2), *expert* (78,6), *Toys“R“Us* (78,2), *Intersport* (77,7), *Mayersche Buchhandlung* (77,1), *Sport 2000* (76,9), *Depot* (74,6), *Thalia* (74,3) sowie *idee+spiel* (73,6). (S. 30) Diese Firmen sind also am stärksten von „Category Killern“ bedroht.

Eingangs ihrer Publikation sprechen die Autoren davon, dass Category Killer das Ende von stationären Händlern auslösen: „Für viele stationäre Buchhändler bedeutete das das Ende. Und auch die CD- und DVD-Abteilungen der Kaufhäuser, Verbrauchermärk-

te und Elektrofachhändler schrumpften merklich zusammen. Doch *Amazon* ist längst nicht der einzige ‚Category Killer‘. Auch in Kategorien, von denen lange behauptet wurde, sie würden sich nie erfolgreich über das Internet vermarkten lassen, gibt es ‚Pure Onliner‘, die der betreffenden Branche das Fürchten lehren.“ (S. 3) Diese und weitere Aussagen sind also so zu verstehen, dass der Kill Thrill-Index die Existenzbedrohung der genannten Branchen und Firmen ausdrückt, zumindest aber schlechtere Erfolge in der Zukunft, wenn sie jetzt nicht gegensteuern.

Solche Aussagen können weitreichende Folgen haben und verschiedene Fragen bei den übrigen Marktteilnehmern aufwerfen. So können sich Kreditgeber fragen, ob die betroffenen Firmen noch in der Lage sind, ihren Kapitaldienst zu leisten, oder genügend Bonität für neue Kredite haben. Kunden sind möglicherweise verunsichert, ob sie im Schadenfall noch ihre Händlergarantie einlösen können oder Dienstleistungen in der Nachkaufphase erhalten, Lieferanten können über die Zahlungsfähigkeit und Einhaltung von Lieferverträgen der genannten Firmen rätseln. Konkurrenten könnten auf frei werdende Standorte spekulieren.

All diese Fragen mögen berechtigt sein, wenn die Aussagen gut begründet und empirisch hinreichend belastbar sind. Dies wollen wir im Folgenden untersuchen. Das Ergebnis nehmen wir schon einmal vorweg: Die Untersuchung ist mangelhaft und ungeeignet, die Schlüsse zu ziehen, wie sie sich in der Publikation finden.

Eine Vorbemerkung: Die Wirtschaftswissenschaften gehören zu den Realwissenschaften. Ihr Erkenntnisanliegen ist es, reale Sachverhalte zu beschreiben, zu erklären und zu prognostizieren. Dies geschieht auf der Grundlage theoretisch gestützter Hypothesen, die im Sinne des kritischen Rationalismus auf ihre Bewährung hin zu überprüfen und im schlechten Fall zu falsifizieren, d. h. zu verwerfen sind. Damit wird eine Basis gelegt, um Gestaltungsempfehlungen zu entwickeln. So gut die theoretische Fundierung und die empirische Untersuchung der wissenschaftlichen Arbeit auch sein mögen, für konkrete Gestaltungsempfehlungen, also Handlungsmaßnahmen für einzelne Firmen, werden die gewonnenen Erkenntnisse allein nicht ausreichen. Es werden immer Informationen über die spezifischen Randbedingungen der Firma benötigt, um die es geht. Der Schluss, deshalb auf Untersuchungen zu verzichten, die wissenschaftlichen Ansprüchen genügen, wäre aber falsch. Denn dann verzichtete man auf vorhandenes Wissen.

An dieser Stelle könnten Bemerkungen kommen, dass Wissenschaftler „reine Theoretiker“ seien und von der Praxis nicht viel verstünden. Solche Bemerkungen sind bekannt, zeigen aber nur, dass diejenigen, die sie äußern, nicht verstanden haben, was „Wissen schaffen“ heißt. Ob sie diese Haltung wider besseres Wissen oder aus Unwissenheit einnehmen, wissen wir nicht. Wir können es nur vermuten. Unwidersprochen dürfte sein, dass es Regeln guter Wissenschaft gibt, die man kennen und anwenden sollte. Dies wollen wir im Folgenden prüfen.

In einem ersten Schritt ist zu prüfen, auf welche Theorie oder theoretischen Aussagen die Studie von *Renz*, *Berentzen* und *Stahlbock* zurückgreift, um vorhandenes Wissen zu nutzen und auf seine empirische Bewährung hin zu überprüfen. Die vorliegende Publikation sagt hierüber nichts aus. Auch implizit, etwa bei der Betrachtung der verwendeten theoretischen Konstrukte – wir gehen hierauf noch ein –, ist nicht zu erkennen, dass eine solche Grundlage vorhanden ist. Zu denken wäre etwa an den Ansatz der Ressourcenabhängigkeit (Resource Dependence Theory) von *Pfeffer* und *Salancik* oder zumindest an einen Untersuchungsrahmen, wie ihn *Porter* mit seinem Modell der Wettbewerbskräfte entwickelt hat. Was ist die Konsequenz, wenn eine solche Grundlage fehlt? Es gibt keine theoretisch fundierten, systematisch hergeleiteten Beziehungen zwischen unabhängigen und abhängigen Untersuchungsgrößen. Alles, was folgt, kann nur unsystematisch und planlos sein.

Sehen wir uns in einem zweiten Schritt an, welche Größen die Autoren untersuchen und wie sie den „Kill Thrill-Index“ ermitteln. Wenn Aussagen darüber gemacht werden sollen, dass bestimmte Bedingungen die Wettbewerbsfähigkeit beeinflussen, dann sind relevante Größen auszuwählen und in Beziehung zueinander zu setzen – dafür wird die Theorie benötigt – sowie zu konzeptualisieren und zu operationalisieren (vgl. z. B. Homburg/Giering 1996). Das Kapitel „Die Methodik der Studie“ sagt lediglich: „Aus der ‚Online-Relevanz der Kategorie‘ und der ‚Online-Performance des Unternehmens‘ setzt sich der Kill Thrill-Index zusammen, der Auskunft über den Bedrohungsgrad einer Branche oder eines Offline-Unternehmens gibt. Dafür sind zwei Größen entscheidend. Erstens, wie hoch die Online-Relevanz der Branche ist (z. B. Anteil Online-Käufer, Relevanz des Internets für das Informations- und Kaufverhalten). Zweitens, wie gut das Unternehmen bereits aufgestellt ist, um das Internet für sich zu nutzen und um sich entsprechend gegen die Bedrohung zu wappnen (Online-Performance des Unternehmens).

Der Kill Thrill-Index ergibt sich aus der Zusammenführung dieser Größen, skaliert auf eine Skala von 0 bis 100.“ (S. 8)

Als Größen für die „Category Online Relevanz“ werden die „Online-Nutzung“, der „Online-Vorteil“, der „Online-Anteil der Kategorie“ und die „Cross Channel-Relevanz“ genannt, für die „Company Online Relevanz“ die Größen die „Online-Nutzung“, der „Online-Kundenanteil“, die „Experteneinschätzung“, die „Cross Channel-Relevanz“ und die „Online-Affinität der Kunden“. (S. 8) Keine dieser Größen wird in der Publikation definiert und operationalisiert. So bleibt z. B. unklar, was unter dem „Online-Vorteil“ und der „Online-Affinität der Kunden“ verstanden wird, wie diese Größen und andere gemessen werden, etwa die „Online-Nutzung“ und der „Online-Kundenanteil“, und wie sich die „Experteneinschätzung“ in den Kanon dieser Größen einreicht. Auch darüber, welche Größen mit wie viel Punkten in den „Kill Thrill-Index“ eingehen, erfährt der Leser an dieser Stelle nichts. Wie immer auch die Punktwerte der einzelnen Größen und in der Summe ausfallen, sie sagen inhaltlich nichts über den Bedrohungsgrad der untersuchten Firmen aus. Denn es mangelt der Untersuchung eben an den erforderlichen Aussagen über den Zusammenhang zwischen erfolgsbeeinflussenden und erfolgsabhängigen Größen. Der Indexwert ist in Bezug auf das Untersuchungsziel schlicht inhaltsleer, d. h. ohne Aussagekraft.

Kommen wir zum dritten Schritt unserer Prüfung, der nun eigentlich überflüssig ist, da die theoretische Basis der empirischen Untersuchung mangelhaft ist, und sehen uns das empirische Design an. Es geht vor allem um die Frage: Welche und wie viele Personen wurden wie zu was befragt? Gehen wir die Punkte im Einzelnen durch. Es gab eine **Kundenbefragung** und eine **Experteneinschätzung**. „Befragt wurden über 8.000 Kunden der betrachteten Händler, die ein repräsentatives Abbild der deutschen Bevölkerung darstellen.“ (S. 7) Statistisch relevant sind z. B. Aussagen darüber, welcher Stichprobenfehler zugrundgelegt worden ist, ob eine geschichtete Stichprobe vorliegt und ob eine Zufallsauswahl oder eine bewusste Auswahl vorgenommen wurde. Darüber findet sich nichts. Dass die befragten Kunden „ein repräsentatives Abbild der deutschen Bevölkerung“ sind, ist eine unbrauchbare Aussage. Erstens kennt der Statistiker nicht den Begriff der Repräsentativität bzw. Repräsentanz, sondern den des Stichprobenfehlers (von der Lippe/Kladroba 2002). Zweitens ist es unsinnig, davon auszugehen, dass die deutsche Bevölkerung die Grundgesamtheit darstelle. Eine solche Aussage

missachtet zum einen, dass nicht alle Einwohner in Deutschland einkaufen können, und zum anderen, dass von denjenigen, die einkaufen können, nicht alle für die Untersuchung relevant sind. Ebenso findet sich nichts über die Anzahl und die Namen der Experten. Der Studie mangelt es hier wie an vielen anderen Stellen an Transparenz, einem zentralen Gütekriterium guter Forschung. Eine Anmerkung: Unsere Bitte an die Autoren, uns das Untersuchungsdesign schriftlich zur Verfügung zu stellen, insbesondere die Fragen, wurde abgelehnt.

Ein weiteres Manko betrifft die **Anzahl der auswertbaren Antworten** pro Firma. Die Autoren schreiben: „Für jedes Unternehmen liegen 150 vollständige Datensätze vor (bis zu drei Unternehmen pro Befragter).“ (S. 7) Wenn sich ein Kunde über „bis zu drei Unternehmen“ (S. 7) geäußert hat, limitiert das die Anzahl der Fragen, die sich beantworten lassen, ohne zu Verzerrungen zu führen. Denn je mehr Fragen zu beantworten sind, um so eher können Ermüdung und Unlust, weitere Fragen zu beantworten, zunehmen und die Reliabilität der Antworten gefährden. Eine größere Anzahl von Fragen ist aber erforderlich, um die vielen Konstrukte ordentlich messen zu können. Insofern lässt die Angabe „Alle erforderlichen Daten und Ergebnisse liegen auf Unternehmensebene vor.“ (S. 32) nicht viel Gehaltvolles erwarten. Fazit: Eine solche Stichprobe ist weit davon entfernt, inhaltlich verwertbare Aussagen zu liefern.

Des Weiteren stellt sich die Frage, was nun genau das **Untersuchungsobjekt** ist, für das Aussagen formuliert werden. Zunächst fällt auf, dass die Eigenschaften der Unternehmen nicht berücksichtigt werden. Wenn wir die genannten Firmen betrachten, so haben wir es mit unterschiedlichen Handelssystemen und jeweils zahlreichen Geschäften zu tun: An Handelssystemen sehen wir erstens Verbundgruppen (z. B. *idee+spiel*, *Intersport*, *Medimax*, *Sport 2000*, *Vedes*), zweitens Filialsysteme (z. B. *Depot*, *Mayersche Buchhandlung*, *Thalia*, *Toys“R“Us*) und drittens hybride Systeme, bestehend aus selbständigen Händlern und Filialen (z. B. *expert*) mit unterschiedlichen Möglichkeiten, ihre Geschäftsmodelle weiterzuentwickeln. *Vedes* ist eine Verbundgruppe mit 1.050 Fachhändlern in Deutschland, Österreich, Italien, Belgien, Luxemburg, Ungarn, den Niederlanden und der Schweiz (<http://bit.ly/1p4DmTS>). *Intersport* hat 1.500 Fachhändler in Deutschland (<http://bit.ly/1IC-C7u7>), *idee+spiel* sind 732 Fachhändler angeschlossenen, davon in Deutschland 692 (<http://bit.ly/SLgiNe>). Zudem haben die Autoren eine Abgrenzung vorgenommen, mit der sie Interdependenzen zerschnei-

den: „Da sich viele Branchen nicht überschneidungsfrei den Kategorien zuordnen lassen, insbesondere die Multisortimenter, erfolgt die Betrachtung des Kill Thrill-Index auf Ebene der Warenkategorien. So lassen sich in Verbindung mit der Online-Performance der Unternehmen individuelle Bedrohungsgrade ermitteln.“ (S. 8) Dies kann man grundsätzlich machen. Dies verlangt dann aber, die Warenkategorien einheitlich abzugrenzen und dann zu erfassen, wie sich die Entwicklung des Online-Handels genau auf diese Warenkategorien auswirkt.

Bezieht sich der „Kill Thrill-Index“ nun auf die gesamte Handelsorganisation, einzelne Händler mit mehreren Geschäften oder einzelne Standorte, d. h. Mitgliedsbetriebe oder Filialen, oder Warengruppen? Darüber erfährt der Leser nichts, und so differenziert sind die Kunden offenbar auch nicht befragt worden oder konnten sich so differenziert nicht äußern. Das Untersuchungsobjekt bleibt daher unscharf, die Zuordnung der Antworten ungeklärt.

Ebenfalls unklar ist, wie die Kunden befragt worden sind. Auch dazu äußern sich die Autoren in dem Kapitel „Die Methodik der Studie“ nicht (S. 7 f.). Wenn es eine Online-Befragung war dann ist die „Online-Affinität“ der Befragten wahrscheinlich höher, als wenn nur oder auch Kunden persönlich, telefonisch oder schriftlich befragt worden wären. Zu vermuten ist außerdem, dass jene Kunden unterrepräsentiert sind, die den Einkauf in stationären Geschäften dem Einkauf im Online-Shop vorziehen.

Unser Fazit: Der Studie mangelt es an der theoretischen Basis ebenso wie an der Einhaltung gängiger Gütekriterien bei der empirischen Erhebung (Transparenz, Validität, Reliabilität). Damit könnte man nun das Kapitel schließen, wäre da nicht noch der Multiplikatoreffekt in den Medien. So greift z. B. die *Wirtschaftswoche* in Heft 14/2014 die Studie „Category Killer“ mit der Titelschlagzeile auf: „Die Online-Opfer – Wo droht der Ladenschluss?“ Und auf den Internetseiten der *Wirtschaftswoche* ist zu lesen: „Online-Angreifer – Diese Unternehmen sind in Gefahr. Immer mehr stationäre Händler geraten wegen der wachsenden Lust der Kunden am Online-Handel in Bedrängnis. Welche Branchen besonders in Gefahr

sind, welche Unternehmen bereits auf der ‚Roten Liste‘ stehen.“ (<http://bit.ly/1ot35Yi>) Angesichts der unübersehbaren Schwächen der Untersuchung ist die unkritische Übernahme der Untersuchungsergebnisse durch die *Wirtschaftswoche* zu bemängeln und kommt einer Rufschädigung der „gerankten“ Unternehmen gleich.

Wir wissen am Ende nicht, ob die Bedrohung der genannten Händler so hoch ist, wie es die Studie darstellt. Es kann so sein, es kann auch anders sein. Die Studie gibt uns dazu keine belastbaren Antworten.

Ob die folgenden Aussagen mehr Gehalt haben, wissen wir auch nicht. Aber man sollte sich damit beschäftigen. So finden sich auf den Internetseiten der *Wirtschaftswoche* Kommentare wie diese: „Spannendes Thema, doch hatte ich hier bei Wi-Wo mehr erwartet. Die Bewertungen erscheinen mir sehr flach und es wiederholen sich immer die Standard-Online-Angreifer.“ (<http://bit.ly/TYRS4i>) Und aus dem Buchhandel kommt die Meldung, dass er „nach langer Talfahrt 2013 um 0,9 Prozent zulegen, während der Online-Handel [mit Büchern, hs] über zwei Prozent verlor.“ (<http://bit.ly/SLgg8c>)

Um nicht missverstanden zu werden: Es geht uns nicht um Sympathien für einzelne Branchen oder Firmen oder um die grundsätzliche Ablehnung von Studien, die nicht aus wissenschaftlichen Einrichtungen kommen. Uns geht es schlicht um die Frage, ob die Regeln guter Wissenschaft eingehalten werden. Nicht mehr, aber auch nicht weniger!

#### Literatur:

Homburg, C.; Giering, A., Konzeptualisierung und Operationalisierung komplexer Konstrukte – Ein Leitfadens für die Marketingforschung, in: *Marketing ZFP*, Heft 1/1996, S. 5-24

Renz, T.; Berentzen, J.; Stahlbock, J.-F., Category Killer – Der stationäre Handel unter Zugzwang, *W&P Fit-Monitor* 2014, Eine Studie von Dr. Wieselhuber & Partner GmbH in Zusammenarbeit mit Facit Research, München 2014

von der Lippe, P.; Kladroba, A., Repräsentativität von Stichproben in: *Marketing ZFP*, Heft 2/2002, S. 227-238 (<http://von-der-lippe.org/dokumente/Repraesentativitaet.pdf>)

Alle Internetseiten sind am 15.6.2014 abgerufen worden.



## Univ.-Prof. Dr. Hendrik Schröder

Hendrik Schröder hat an der Westfälischen Wilhelms-Universität Münster Betriebswirtschaftslehre studiert, dort promoviert und sich im Fach Betriebswirtschaftslehre habilitiert.

Seit 1996 ist er Inhaber des Lehrstuhls für Betriebswirtschaftslehre, insbesondere Marketing und Handel, an der Universität Duisburg-Essen ([www.marketing.wiwi.uni-due.de](http://www.marketing.wiwi.uni-due.de)).

Zudem ist Professor Schröder Leiter des Forschungszentrums für Category Management in Essen ([www.cm-net.wiwi.uni-due.de](http://www.cm-net.wiwi.uni-due.de)) und Direktor des Internationalen Centruns für Franchising und Cooperation ([www.franchising-und-cooperation.de](http://www.franchising-und-cooperation.de)).

## Lehrstuhl für Marketing und Handel

Universität Duisburg-Essen, Campus Essen  
Fachbereich Wirtschaftswissenschaften  
Universitätsstr. 12  
45141 Essen  
t +49 (201) 183 - 3385  
f +49 (201) 183 - 2833  
[www.marketing.wiwi.uni-due.de](http://www.marketing.wiwi.uni-due.de)  
[hendrik.schroeder@uni-duisburg-essen.de](mailto:hendrik.schroeder@uni-duisburg-essen.de)

ISSN 2364-4443 (Print)  
ISSN 2364-4451 (Internet)