

Andreas Oehler

Testen der Tester? Grundsätze ordnungsgemäßen Testens!

Tests und Vergleiche von Gütern und Dienstleistungen sind mittlerweile weit verbreitet. Würde der einzelne Verbraucher den Vergleich selbst durchführen, wäre dies mit hohen Kosten verbunden, institutionelle Tests bieten demgegenüber eine gute Informationsmöglichkeit. Die Vertrauenswürdigkeit der Tester ist dabei von herausragender Bedeutung. Der Autor stellt Grundsätze für ordnungsgemäßes Testen auf und schlägt vor, wie diese kontrolliert werden können.

Tests und ihre vergleichenden Ergebnisse können für Verbraucher, aber auch als Fremdeinschätzung für seriöse Anbieter, sowie für die Verbraucherpolitik (Marktwächter) wertvolle Informationen zur Qualität und Preiswürdigkeit von Sachgütern und Dienstleistungen liefern. Die damit verbundene Orientierung für die Marktteilnehmer in einer keineswegs voll transparenten und funktionierenden Marktwirtschaft wie in Deutschland erlaubt eine Verbesserung des Marktgeschehens. Insofern agieren Tester als Informationsintermediäre, die grundsätzlich die Möglichkeit schaffen, eine arbeitsteilige Wirtschaft und Gesellschaft zu unterstützen, da nicht jeder alles selbst untersuchen, probieren und vergleichen kann und muss. Auf einer Meta-Informationsebene wohnt Tests und Vergleichen die Chance inne, dem Urteil Dritter zu folgen, ohne jedes Vertrauen hinterfragen und aufwendige Recherchen selbst durchführen zu müssen, soweit dies überhaupt möglich ist. Damit dies gelingen kann, bedarf es einiger weniger Grundsätze oder Mindestanforderungen.

Orientierung für potenzielle Nachfrager

Das Testen und Vergleichen von Sachgütern und Dienstleistungen ist eine inzwischen weit verbreitete Praxis auf nahezu allen Märkten in Deutschland, an denen Verbraucher teilnehmen und mitwirken. Tests und Vergleiche sollen meist die Qualität und gegebenenfalls auch die Preiswür-

digkeit der untersuchten Sachgüter und Dienstleistungen bezogen auf die möglichen Produkteigenschaften und Nutzungsmöglichkeiten zusammenfassen und insbesondere eine Orientierung für die potenziellen Nachfrager auf dem Markt, aber gegebenenfalls auch eine Fremdeinschätzung für die konkurrierenden Anbieter erlauben.

Gut bekannt und weiter verbreitet sind die Untersuchungen der Stiftung Warentest, der Verbraucherzentralen, des Verbraucherservice oder von Öko-Test. Nimmt man die ganze Bandbreite der sehr unterschiedlich angelegten Untersuchungen von Verbrauchersendungen in Radio und Fernsehen, z.B. ARD-Ratgeber, Frontal, Klartext, markt, Markencheck, Marktcheck, mEx, nano, Plusminus, Panorama, Report, Servicezeit, WiSo etc., oder von Nicht-Regierungsorganisationen wie foodwatch, WWF, NABU, BUND, Greenpeace etc. oder von Vergleichsportalen für Sachgüter und (Finanz-)Dienstleistungen (Versicherungen, Immobilienkredite etc.) bis hin zur Energieversorgung (Strom, Gas etc.) sowie schließlich Anbieter wie z.B. ADAC, TÜV dazu, dann kann man wohl längst von einer Testindustrie sprechen.

Die folgenden Überlegungen gehen von der scheinbar einfachen Fragestellung aus, welche Mindestanforderungen für eine gute Testpraxis gelten sollten, um eine valide Mindestqualität der (vergleichenden) Testurteile zu gewährleisten. Ein solcher Mindeststandard ließe sich als „Grundsätze ordnungsgemäßen Testens“ (GoT) oder mit einer Reminiszenz an das anglo-amerikanische Mutterland des vergleichenden Testens als „Generally Accepted Testing Principles“ (GATP) kennzeichnen. Erst auf diese Weise würde die genannte Orientierung für Verbraucher wie auch Anbieter zur Qualität und Preiswürdigkeit von Sachgütern und Dienstleistungen valide möglich.

Die nachfolgend erörterten Standards sind nicht einfach nur als „best practice“ zu verstehen, sondern sie leiten sich ebenso aus theoretischen wie empirischen Analy-

Prof. Dr. Andreas Oehler ist Lehrstuhlinhaber für Betriebswirtschaftslehre, insbesondere Finanzwirtschaft sowie Direktor der Forschungsstelle Verbraucherfinanzen & Verbraucherbildung an der Universität Bamberg.

sen der verbraucherwissenschaftlichen Forschung,¹ z.B. aus der Institutionenökonomie und der Verhaltensökonomie, ab.² Ohne auf die vielfältigen Studien an dieser Stelle im Einzelnen eingehen zu können, gehören hierzu z.B. die Erkenntnisse, dass die bekannten Informationsasymmetrien zwischen Anbietern und Nachfragern nicht vollständig eliminierbar, sondern als systemimmanent zu bezeichnen sind, insbesondere bei Erfahrungs- und Vertrauensgütern. Ebenso führen regelmäßige „biases“ der Informationswahrnehmung und -verarbeitung bei allen Marktbeteiligten zu einer ganz normalen Entscheidungsüberlastung und Orientierungshürden.

Mindestanforderungen für Tests von Gütern und Dienstleistungen

Mindestanforderungen sind geeignet, vorhandene Testsysteme zu überprüfen, auch durch eine staatliche Aufsicht, aber ebenso, um bei der Entwicklung neuer Strukturen zu helfen.³ Drei Grundsätze lassen sich hier unterscheiden:

Grundsatz 1 – Identifizierung und Transparenz

Ist klar erkennbar, wer das Testurteil vergibt? Gibt sich die dahinter stehende Organisation oder Institution klar zu erkennen? Wer finanziert einen Test oder Vergleich, in welcher wirtschaftlichen, insbesondere finanziellen Beziehung stehen Tester, Anbieter oder Testnutzer?

Während Anzeigen von Anbietern in Print- und Online-Medien es auf den ersten Blick zu erlauben scheinen, Financiers zu erkennen, so schwer wird es auf den zweiten Blick, deren tatsächliche Relevanz für die Tests selbst abzuleiten. Gerade die Nicht-Berücksichtigung von Sachgütern oder Dienstleistungen einzelner Anbieter in Untersuchungen, oder die nicht gegebenen Informationen machen es außenstehenden Nutzern nahezu unmöglich, ein mögliches Interessengeflecht zu erkennen. Insofern kann man allein schon

- 1 Vgl. für einen Überblick A. Oehler, L. Reisch: Sie lebt! Zur Verbraucherforschung im deutschsprachigen Raum: Eine empirische Analyse, in: Journal für Verbraucherschutz und Lebensmittelsicherheit, Bd. 7 (2012), H. 2, S. 105-115, und die dort zitierte Literatur.
- 2 Vgl. für einen Überblick A. Oehler: Behavioral Economics und Verbraucherpolitik: Grundsätzliche Überlegungen und Praxisbeispiele aus dem Bereich Verbraucherfinanzen, in: BankArchiv, 59. Jg. (2011), H. 10, S. 707-727; ders.: Klar, einfach, verständlich und vergleichbar: Chancen für eine standardisierte Produktinformation für alle Finanzdienstleistungen. Eine empirische Analyse, in: ZBB Zeitschrift für Bankrecht und Bankwirtschaft, 24. Jg. (2012), H. 2, S. 119-133; ders.: Neue alte Verbraucherleitbilder: Basis für die Verbraucherbildung?, in: HiBiFo Haushalt in Bildung und Forschung, 2. Jg. (2013), Nr. 2, S. 44-60.
- 3 Vgl. A. Oehler: Labeling & Co – Wegweiser, Irreführung oder Wettbewerbsanreiz?, in: EMF Spectrum, 4. Jg. (2013), Nr. 3, S. 31-34. Bezogen auf die Verwendung von Vertrauenslabels hat die Verbraucherkommission Baden-Württemberg bereits 2011 einen Forderungskatalog erstellt, der im Folgenden ebenfalls berücksichtigt wird. Vgl. VK BW: Vom Labelmissbrauch zu Vertrauenslabels, Stellungnahme, Stuttgart 2011.

hier die deutliche Wechselwirkung zu den Grundsätzen 2 und 3 erkennen: Von der Anlage der Untersuchung, vom Testdesign oder auch von der Marktrelevanz der Tests und Vergleiche hängt wesentlich ab, ob überhaupt eine unabhängige Aussage zu erwarten ist. Andererseits ist allein die Vollständigkeit einer Untersuchung nur begrenzt aussagefähig, wenn die getesteten Produkte vor Ort für Verbraucher nicht erhältlich sind (Beispiele: Dispozinsen, Bankgebühren).

Grundsatz 2 – Verifizierung

Sind die Kriterien und Prüfmethode öffentlich zugänglich und durch Dritte nachprüfbar? Wird offengelegt, wie geprüft wird und welche Standards für welche Untersuchung zugrundegelegt werden?

Zumindest gegenüber den Gremien mit Aufsichtsfunktion und weiteren Sachverständigen muss eine Offenlegung der wesentlichen Merkmale eines Testdesigns samt der zugehörigen Vorüberlegungen und Auswahlprozesse frühzeitig möglich sein, nicht erst nach Durchführung und Publikation. Aufgrund der erheblichen Bedeutung von Tests und Vergleichen für die Entscheidungen von Verbrauchern wie Anbietern sollte auf diesen Grundsatz der Verifizierung ein hohes Gewicht gelegt werden, um rechtzeitig fehlleitende Informationen aus Untersuchungen zu vermeiden. Dem dient auch der Schärfung des Bewusstseins für die Bedeutung von Testinformationen auf Märkten, ein Schutz vor krimineller Energie wird jedoch nur bedingt möglich sein. Möglich erscheint aber eine frühzeitige Klärung, inwieweit Testurteile aus Indizien und Einschätzungen oder entlang wissenschaftlicher Standards abgeleitet und nachweisbar sind.

Grundsatz 3 – Relevanz und Nützlichkeit

Sind die Aussagen eines Tests oder Vergleichs maßgeblich für die wesentlichen Eigenschaften eines Sachguts oder einer Dienstleistung? Wird der Fokus nur auf einzelne Teilaspekte gelegt, z.B. sozial-ökologische Kriterien, oder werden alle wesentlichen Eigenschaften gleichermaßen berücksichtigt? Gehen die Test- und Vergleichskriterien über staatliche Standards hinaus oder überprüfen sie im Wesentlichen nur deren Einhaltung?

Die genannten Kriterien beziehen sich zunächst darauf, sicherzustellen, dass tatsächlich und gegenständlich auch getestet wird, was angekündigt wird: Ist im Test „drin“, was auch „drauf“ steht? So relevant für manche vielleicht die Geräuschemission einer Waschmaschine sein mag, Kern ist zunächst die Waschleistung und gegebenenfalls der Ressourcenverbrauch. Wesentlich ist hierbei auch, um möglichst viele, unterschiedliche Verbraucher effizient zu

erreichen, dass die Gewichtungsfaktoren der Testkriterien gut sichtbar und verständlich offengelegt werden, damit jeder Nutzer gegebenenfalls eigene Beurteilungen vornehmen kann. Unklare Verrechnungen und „scorings“ behindern die Funktionalität von Testinformationen.

Weitere Kriterien gehören zum Grundsatz der Relevanz und Nützlichkeit: Steht die Verwendung der Testergebnisse und die Nachvollziehbarkeit der Untersuchung für Verbraucher oder einzelne Gruppen im Zentrum oder werden eher allgemeine Aussagen ohne konkrete Handlungsempfehlungen getroffen? Werden Sachgüter und Dienstleistungen getestet, die viele Verbraucher betreffen oder eher Marktnischen? Werden Lösungsvorschläge und Handlungsempfehlungen abgeleitet, die an konkreten Bedarfssituationen der Verbraucher ausgerichtet sind, statt alleine auf Anbieter bzw. Produkte?

Hier sollte für jeden Nutzer schnell, einfach und verständlich erkennbar sein, ob ihm der Test oder der Vergleich überhaupt einen Nutzen stiftet. Wer z.B. zur weit überwiegenden Mehrheit der deutschen Anleger gehört, der – wie oft sinnvoll und empfohlen – für die mittel- und langfristige Geldanlage gar keine aktiv gemanagten Investmentfonds nutzt, der sollte auf einen Blick erkennen können, dass eine Zeitungsbeilage am Wochenende oder ein Fondstest in einer Zeitschrift oder online für ihn keinen Nutzen stiftet. Wer sich beispielsweise für „Riester-Produkte“ interessiert, der wird klare Angaben, die alle Varianten direkt vergleichen, sehr schätzen. Verbraucher können dann die vielen unproduktiven reinen produkt- und anbietergetriebenen Einzeltests schnell ignorieren. Gleichzeitig muss direkt sichtbar sein, ob Handlungsempfehlungen und Lösungsvorschläge wesentliche Bedarfe und Produkt-Komponenten berücksichtigen (z.B. alle Kosten; beispielsweise fehlen bei den „Riester-Vergleichen“ oft die Verrentungskosten bei Banksparrplänen, hier führt ein Test eher in die Irre). Mit anderen Worten: Für Tests und Vergleiche müssen Mindestanforderungen der Kennzeichnung gelten.

Wirtschaftliche Bedeutung der Testindustrie

Aufgrund der inzwischen erreichten erheblichen wirtschaftlichen Bedeutung der Testindustrie und der grundsätzlich vorhandenen maßgeblichen Wirkung auf vielfältige Entscheidungen von Verbrauchern, aber auch auf das Angebot und die Beschaffenheit von Produkten durch die anbietende Wirtschaft, wird inzwischen gerne ein „Testen der Tester“ gefordert. So verständlich solche Überlegungen auf den ersten Blick erscheinen, so zerrinnen diese Ideen bei näherer Betrachtung: Solcherlei Vorgehen macht nämlich Vertrauen nicht überflüssig, sondern tauscht nur die Zielgruppen aus, auf die sich das Vertrauen richten soll. Es

deutet sich ein infinites Regress an, die Kontrolle der Kontrolle der Kontrolle ...

Basisregulierung erforderlich

Abhilfe kann aber eine inhaltliche und formale Überwachung (Marktaufsicht) auf der Basis der genannten Grundsätze als Sicherstellung einer Mindestqualität schaffen, also eine Basisregulierung. Hierzu gehören bezogen auf die drei Grundsätze der Nachweis eines internen Risikomanagements mit Kontroll- und Fehlermanagementsystemen und Gremien, die die betroffenen Verbraucher und Anbieter ebenso einbeziehen wie unabhängige wissenschaftliche Expertise und sachverständige Personen, z.B. unabhängiger Medien.

Die Idee und Zielsetzung einer Basisregulierung sollte sich aufgrund der Vor- und Nachteile regulierender Eingriffe auf die genannten Grundsätze (GoT/GATP) im Sinne eines Mindeststandards beschränken. Damit würden mögliche Nachteile einer weitergehenden Regulierung wie die Beschränkung des Testens nach Testproduktauswahl, Testverfahren oder Testpublikation samt damit verbundenem bürokratischem Aufwand vermieden oder gering gehalten. Gleichzeitig blieben die Vorteile der Mindestregulierung erhalten. Hierzu gehören unter anderem die Gewährleistung der Orientierung für Verbraucher und Anbieter, die Sicherstellung des Markteintritts neuer Wettbewerber/Tester, die Möglichkeit für alternative Testverfahren und Testpublikationen und eine grundsätzlich von Angebot und Nachfrage abhängige Auswahl der zu testenden Sachgüter und Dienstleistungen.

Regelmäßige Überprüfung durch Wirtschaftsprüfer

Die Einhaltung des rechtlich verbindlichen Mindeststandards der Grundsätze ordnungsgemäßen Testens benötigt kein aufwendiges Aufsichtswesen, insbesondere kann auf bestehende Regelungen in verwandten Themenfeldern rekurriert werden. Inhaltlich muss über die gesamte Testprozesskette ein „gelebtes“ Risikomanagement- und Qualitätssicherungssystem nachgewiesen werden. Die regelmäßige Überprüfung kann z.B. im Rahmen der Jahresabschlussprüfung durch die bestellten Wirtschaftsprüfer mit entspre-

chendem Prüfauftrag vorgenommen werden, so wie dies heute schon für andere Themenbereiche üblich ist.

Zu diesem „Lebendigkeits-Nachweis“ für das implementierte interne Risikomanagement mit Kontroll- und Fehlermanagementsystemen gehört auch die Zusammensetzung und Arbeit der eingesetzten Gremien und die notwendige Unabhängigkeit der beauftragten Personen, z.B. in Fachgremien, die das Testverfahren erörtern und begleiten, oder in Gremien, die für die Initiierung von Testvorhaben zuständig sind. Einbezogen werden sollten auch die installierten oder zu installierenden Aufsichtsorgane der Geschäftsführung der Testinstitutionen, wie z.B. Verwaltungsrat oder Beirat, die entsprechende Informations- und Mitwirkungsrechte brauchen. Entscheidend sollte dabei sein, dass die eingerichteten internen Gremien und Prozesse von der Festlegung eines Testvorhabens und der extern zu Rate gezogenen Expertise (z.B. Prüflabore, Marktforscher) über die „Probenziehung“ bis hin zur journalistischen sowie rechtlichen Bewertung vor möglichen Publikationen volle Funktionsfähigkeit zeigen.

Aufgrund der angesprochenen Entwicklung der letzten Jahre in Richtung einer regelrechten Testindustrie wäre es zudem naheliegend, eine regelmäßige Marktüberwachung im Sinne einer Vermeidung enger Oligopole und ihrer Negativanreize zu initiieren, angesiedelt in den bisherigen Überwachungsstrukturen als ergänzendes Themenfeld des Bundeskartellamtes.

Das Zusammenwirken der Beteiligten gibt die Chance, das Vertrauen in Testurteile zu rechtfertigen und einem weiteren Vertrauensverlust durch wirkungslose Kontrollpyramiden entgegenzuwirken (Aufseher der Aufseher der Aufseher ...). Die vorgeschlagenen Grundsätze ordnungsgemäßen Testens sichern den wichtigen Informationscharakter vergleichender Tests und stärken die notwendige Reputation. Es liegt im Interesse der Verbraucher, der Anbieter und seriöser Tester, dauerhaft sorgfältig recherchierte, überprüfbare, relevante und nachvollziehbare Testergebnisse zu publizieren und nicht mit dem Fokus auf schnelle Schlagzeilen das wertvolle Gut verlässlicher Testinformationen zu beschädigen.

Title: *Testing the Testers? Generally Accepted Testing Principles!*

Abstract: *Tests and comparisons of products can provide useful information on both quality and value of goods and services for consumers, suppliers, and for consumer policy, for example watchdogs. In a market economy like Germany, where consumer markets are not fully transparent and functioning, testing organisations act as information intermediaries. They provide information at the meta-level so that not everyone has to test and compare all goods and services individually. This is why the reliability of testers and testing organisations is of particular importance. This article examines the principles or minimum standards for proper testing and discusses how to control them.*

JEL Classification: D11, D18, K2