



Bundeskartellamt

Sektoruntersuchung Nutzerbewertungen

Bericht

Oktober 2020

Sektoruntersuchung Nutzerbewertungen

Bericht gemäß § 32e GWB

Az. V-22/19

Oktober 2020

Kontakt

Bundeskartellamt

Beschlussabteilung Wettbewerbs- und Verbraucherschutz

Kaiser-Friedrich-Straße 16

53113 Bonn

poststelle@bundeskartellamt.bund.de

www.bundeskartellamt.de

Inhaltsverzeichnis

A. Zusammenfassung	1
B. Einleitung	4
B.1. Anlass und Ziele der Sektoruntersuchung Nutzerbewertungen	4
B.2. Gegenstand der Sektoruntersuchung Nutzerbewertungen	6
C. Ermittlungen.....	8
C.1. Befragung von Portalen, die Nutzerbewertungen anzeigen.....	8
C.1.1. Auswahl der zu befragenden Portale	9
C.1.2. Durchführung der Befragung.....	10
C.2. Stellungnahmen von Bewertungsvermittlern.....	11
C.3. Gespräche mit Marktteilnehmern und weiteren Institutionen	12
C.4. Konsultation	12
D. Funktionsweisen von Bewertungssystemen	14
D.1. Erfassung von Bewertungen	15
D.1.1. Welche Leistungen können bewertet werden?	15
D.1.2. Wer kann Bewertungen abgeben?.....	16
D.1.3. Anstupser und Anreize zur Bewertungsabgabe	17
D.1.4. Produkttests	19
D.1.5. Wesentliche Erkenntnisse	20
D.2. Filterung der Bewertungen	21
D.2.1. Automatisiertes Filtern.....	22
D.2.2. Manuelles Filtern	25
D.2.3. Nutzerinitiierte nachträgliche Überprüfungen.....	26
D.2.4. Anbieterinitiierte nachträgliche Überprüfungen.....	28
D.2.5. Konsequenzen und Sanktionen	32
D.2.6. Wesentliche Erkenntnisse	35
D.3. Darstellung der Bewertungen	37
D.3.1. Einzeldarstellung	38
D.3.2. Zusammenfassende Darstellung	41
D.3.3. Wesentliche Erkenntnisse	43
E. Interessenlagen und Verhaltensweisen der Marktteilnehmer.....	45
E.1. Verbraucher	45
E.1.1. Verbraucher als Leser von Nutzerbewertungen.....	47
E.1.2. Verbraucher als Verfasser von Nutzerbewertungen	47
E.1.3. Wesentliche Erkenntnisse	52

E.2.	Anbieter.....	52
E.2.1.	Einfluss von Bewertungen auf den Umsatz der Anbieter.....	53
E.2.2.	Generierung von Nutzerbewertungen durch Anbieter	55
E.2.3.	Wesentliche Erkenntnisse	59
E.3.	Bewertungsvermittler	59
E.3.1.	Vermittlung von Bewertungen über Unternehmen.....	61
E.3.2.	Vermittlung von Bewertungen über Gruppen in (sozialen) Netzwerken.....	64
E.3.3.	Wesentliche Erkenntnisse	66
E.4.	Portale	67
E.4.1.	Einfluss von Bewertungen auf den Umsatz der Portale	68
E.4.2.	Politische und gesetzgeberische Initiativen in Bezug auf Portale	71
E.4.3.	Wesentliche Erkenntnisse	73
F.	Problemdifferenzierung und Lösungsansätze	75
F.1.	Incentivierte Bewertungen	77
F.1.1.	Ökonomische und rechtliche Bewertung.....	77
F.1.2.	Lösungsansatz: Bedarfsgerechter Einsatz von gekennzeichneten Anreizen.....	79
F.2.	Produkttest-Bewertungen	81
F.2.1.	Ökonomische und rechtliche Bewertung.....	81
F.2.2.	Lösungsansatz 1: Stärkere Inpflichtnahme der Portale.....	84
F.2.3.	Lösungsansatz 2: Kennzeichnung von Produkttest-Bewertungen	88
F.3.	Manipulierte Bewertungen.....	91
F.3.1.	Ökonomische und rechtliche Bewertung.....	91
F.3.2.	Lösungsansatz: Effektivere Rechtsdurchsetzung.....	93
F.4.	Nicht-authentische Bewertungen	93
F.4.1.	Ökonomische und rechtliche Bewertung.....	94
F.4.2.	Lösungsansatz: Stärkere Inpflichtnahme der Portale.....	96
F.5.	Asymmetrisch selektierte Bewertungen.....	97
F.5.1.	Ökonomische und rechtliche Bewertung.....	98
F.5.2.	Lösungsansatz: Beseitigung asymmetrischer Komponenten	100
F.6.	Weitere Problembereiche.....	101
F.6.1.	Unklare Herkunft und Auswahl von Kundenbewertungen	101
F.6.2.	Missverständliche Durchschnittsangaben.....	102
F.6.3.	Fehlende Repräsentativität	102
G.	Fazit.....	104
H.	Anhang.....	106

A. Zusammenfassung

Nutzerbewertungen im Internet sind für Verbraucher unbestritten von großer Bedeutung. Wenn Bewertungen allerdings nicht in der erwarteten Weise zustande kommen, können sie Verbraucher dazu verleiten, falsche Entscheidungen zu treffen. Dies gilt insbesondere für systematische Verzerrungen, die Verbraucher kaum erkennen können. Seit einigen Jahren ist das Problem unter den Schlagworten „Fake-Bewertungen“ und „gefälschte Bewertungen“ immer wieder Gegenstand zahlreicher Studien, Verfahren und Berichte, die sich jedoch häufig in pauschalen Verurteilungen oder in Einzelfallbetrachtungen aus Unternehmenssicht erschöpfen. Ziel der vorliegenden Sektoruntersuchung ist es, eine differenzierte Gesamtbetrachtung des Themas Nutzerbewertungen vorzunehmen, um Verbraucher, Wirtschaft, Politik und Wissenschaft besser über die Hintergründe und die Zusammenhänge zu informieren, mutmaßliche Verbraucherrechtsprobleme aufzuklären und nicht zuletzt mittels konkreter Lösungsansätze den Verbraucherschutz in diesem Bereich zu verbessern. Das Bundeskartellamt hat dazu die Funktionsweisen von Bewertungssystemen untersucht, die Interessenlagen der verschiedenen Marktteilnehmer analysiert und eine Differenzierung der in der Praxis relevanten Problembereiche vorgenommen. ► Kapitel B.

Im Rahmen der Ermittlungen wurde in erster Linie eine unmittelbare und umfangreiche schriftliche Befragung von 66 Portalen durchgeführt, die Nutzerbewertungen für Produkte, Unternehmen oder Apps aus insgesamt 16 verschiedenen Branchen anzeigen. Bei den ausgewählten Portalen handelt es sich um große Plattformen, Bewertungsportale, Suchmaschinen, Online-Shops etc., die für Verbraucher in Deutschland besonders relevant sind. Darüber hinaus hat das Bundeskartellamt verschiedene Unternehmen, die die Vermittlung von Bewertungen oder Produkttests anbieten, um Stellungnahme gebeten sowie eine Vielzahl von Gesprächen mit weiteren Marktteilnehmern, Verbänden und Behörden geführt. Ergänzt wurden die Ermittlungen durch die Auswertung zahlreicher Studien und Gerichtsverfahren. Nach Veröffentlichung des Konsultationspapiers wurden zudem die von den Marktteilnehmer dazu eingereichten Stellungnahmen berücksichtigt. ► Kapitel C.

Die umfassenden Ermittlungen zu den von den Portalen verwendeten Bewertungssystemen haben in Bezug auf die Erfassung, die Filterung und die Darstellung von Nutzerbewertungen detaillierte und differenzierte Erkenntnisse erbracht: Bei der **Erfassung** von Nutzerbewertungen gibt es große Unterschiede bzgl. der Voraussetzungen für die Aufnahme eines zu bewertenden Produkts oder Unternehmens und die Abgabe einer Bewertung sowie den Einsatz von Anstupsern, Anreizen oder Produkttests zur Generierung von Bewertungen. Insbesondere asymmetrische Anstupser können hier die Repräsentativität der abgegebenen Bewertungen beeinträchtigen. Bei der **Filterung** der abgegebenen Bewertungen nehmen viele Portale eine automatische oder manuelle Vorab-Prüfung auf Schimpfworte, Werbung oder Datenschutzverstöße vor. Nur wenige Portale setzen darüber hinaus auch komplexere Filter ein, die notwendig sind, um nicht-authentische oder manipulierte Bewertungen zu identifizieren, und sanktionieren systematisch Verstöße. Andere verlassen sich eher auf nachträgliche Meldungen von Nutzern und Anbietern. Hinsichtlich der **Darstellung** von Nutzerbewertungen haben die Ermittlungen teilweise in Bezug auf die Basis oder die Berechnung der angezeigten Durchschnittsbewertungen Probleme aufgedeckt. Für alle drei Komponenten Erfassung, Filterung und Darstellung gilt gleichermaßen, dass die Funktionsweise des jeweils verwendeten Bewertungssystems für den Verbraucher transparent sein sollte. Dabei ist jedoch zu berücksichtigen, dass die Aufnahmefähigkeit und die

Aufnahmebereitschaft des Verbrauchers auch oder gerade im Internet begrenzt sind, sodass ein Mehr an Informationen nicht automatisch zu einem Mehr an Transparenz führt. ► Kapitel D.

Um das Thema Nutzerbewertungen in seiner gesamten Komplexität erfassen und analysieren zu können, ist die Betrachtung der damit verbundenen Verhaltensweisen und Interessenlagen der verschiedenen Marktteilnehmer unerlässlich: So wollen die meisten **Verbraucher** zwar viele und hilfreiche Bewertungen lesen, nur wenige einzelne Verbraucher sind jedoch in bestimmten Situationen dazu bereit, selbst eine Bewertung zu schreiben. Das führt u. a. zum bislang kaum beachteten Problem der mangelnden Repräsentativität. Für die **Anbieter** bedeutet eine große Zahl an Bewertungen, dass ihre eigenen Produkte bzw. Dienstleistungen von den Nutzern besser gefunden und eher ausgewählt werden. Die Anbieter können das Bewertungsverhalten der Verbraucher aber nur sehr bedingt beeinflussen und unterliegen zudem den Bewertungsrichtlinien der Portale, sodass sie teilweise auf die Vermittlung von Bewertungen und Produkttests durch Dritte zurückgreifen. Aufgrund der Tatsache, dass inzwischen fast alle Portale Bewertungen durch Produkttester auf ihrer Plattform unterbinden, können sich **Bewertungsvermittler**, die authentische und nicht-manipulierte Bewertungen generieren lassen, gegenüber ihren weniger seriösen Konkurrenten und deren Garantie auf Top-Bewertungen allerdings kaum mehr behaupten. Die **Portale** selbst schließlich haben grundsätzlich ebenfalls ein hohes Interesse an vielen Bewertungen und wollen deren Glaubwürdigkeit sicherstellen, werden aber teilweise erst bei offensichtlichen Problemen aktiv, weil sie den Aufwand einer sorgfältigen Prüfung von Bewertungen scheuen und möglicherweise auch nicht die Verantwortung für irreführende Bewertungen übernehmen möchten. ► Kapitel E.

Auf der Grundlage dieser Erkenntnisse können zu den bestehenden Problemen einzelne Kategorien gebildet werden, die jeweils eine differenzierte Betrachtung erfordern und für die sich differenzierte Lösungsansätze ergeben. So können **Incentivierungen und Produkttests** grundsätzlich dazu beitragen, mehr hilfreiche Bewertungen zu generieren, wenn diese für den Verbraucher klar gekennzeichnet sind und ihr Zustandekommen transparent ist; ansonsten sind sie lauterkeitsrechtlich unzulässig. **Manipulierte oder nicht-authentische Bewertungen** hingegen sind ökonomisch schädlich und lauterkeitsrechtlich unzulässig, sodass Portale für deren Identifizierung und Entfernung stärker in die Pflicht genommen werden sollten. Direkt verantwortlich sind Portale und ggfs. Anbieter, wenn sie bei der Abgabe oder Filterung von Bewertungen einseitig negative Bewertungen benachteiligen (**Asymmetrie**). Entsprechende Mechanismen sind unzulässig und müssen beseitigt werden. ► Kapitel F.

Um zu vermeiden, dass Verbraucher durch verschiedene Arten gefälschter Bewertungen oder Bewertungssysteme getäuscht werden und damit möglicherweise auch insgesamt das Vertrauen in Nutzerbewertungen verlieren, sind verschiedene Maßnahmen vorstellbar. Alle Portale müssen Filter nutzen, mit denen sie Bewertungen vor der Veröffentlichung auf Authentizität und mögliche Manipulationen überprüfen. Sie können monetäre Anreize zur Abgabe von Bewertungen setzen, sofern sie diese entsprechend kennzeichnen. Manche Portale sind in diesem Bereich bereits sehr aktiv, viele Portale schöpfen aber die diesbezüglich bestehenden Möglichkeiten bei Weitem nicht aus. Handelsportale sollten bei Bedarf die Einstellung gekennzeichnete Produkttest-Bewertungen auf ihren Seiten zulassen. Mangels behördlicher Zuständigkeiten zur Durchsetzung des geltenden Verbraucherrechts sollte zudem die zivilrechtliche Verfolgung von Rechtsverstößen bei Nut-

zurbewertungen – im Rahmen ihrer Nachweismöglichkeiten – in Fortentwicklung der bestehenden Rechtsprechung auf eine klarere und stärkere Verantwortlichkeit der Portale für die auf ihren Webseiten begangenen Verbraucherverstöße dringen. ► Kapitel G.

Hinweise für Verbraucher

Was sind die wesentlichen Risiken bei Nutzerbewertungen?

- Bewertungen im Internet stammen teilweise gar nicht von echten Nutzern oder werden inhaltlich manipuliert. Dies kann z. B. bei Bewertungen von Ärzten, von Hotels oder auch von Produkten der Fall sein und ist für Sie als Verbraucher oft nicht erkennbar.
- Manche Bewertungen kommen nur deshalb zustande, weil der Verfasser hierfür einen Anreiz erhalten hat, z. B. die kostenlose Bereitstellung des betreffenden Produkts. Dieser Anreiz kann dazu führen, dass die Bewertung tendenziell etwas positiver ausfällt.
- Auf einigen Portalen werden die von Nutzern abgegebenen positiven Bewertungen eher dargestellt als die negativen Bewertungen, z. B. weil die Portale negative Bewertungen häufiger entfernen (müssen), nachdem die betroffenen Anbieter wie Ärzte oder Restaurants eine Bewertung als kritisch gemeldet haben. Die daraus resultierende Verzerrung des Gesamtbilds der Bewertungen hin zum Positiven ist für Sie als Verbraucher in der Regel nicht zu erkennen.

Worauf sollte man beim Lesen von Nutzerbewertungen achten?

- Bedenken Sie, dass Nutzerbewertungen auf unterschiedliche Weise entstehen können und behalten Sie die oben genannten Risiken im Hinterkopf. Achten Sie insbesondere auf Anzeichen dafür, dass eine Bewertung nicht-authentisch oder manipuliert sein könnte.
- Lesen Sie falls möglich immer mehrere Bewertungen zu einem Produkt bzw. Unternehmen. Sehen Sie sich im Zweifel zudem andere Bewertungen desselben Nutzers an, um dessen Glaubwürdigkeit zu überprüfen.
- Nutzen Sie ggfs. die von einigen Portalen angebotene Möglichkeit, auffällige Bewertungen zu melden oder als kritisch zu markieren.
- Falls Sie mehr über das Zustandekommen der Nutzerbewertungen auf einem Portal wissen wollen, machen Sie sich mit den Richtlinien vertraut, in denen das Portal seine Bedingungen und Maßnahmen in Bezug auf Bewertungen erläutert.
- Geben Sie auch solchen Produkten oder Dienstleistungen eine Chance, die noch keine oder nur wenige Bewertungen haben. Eine große Zahl von Bewertungen ist nicht automatisch ein Qualitätskriterium.

B. Einleitung

B.1. Anlass und Ziele der Sektoruntersuchung Nutzerbewertungen

Verbraucher, die im Internet nach Produkten oder Dienstleistungen suchen, orientieren sich in starkem Maße an den dort angezeigten Kunden- bzw. Nutzerbewertungen. Die von anderen Nutzern abgegebenen Bewertungen in Form von Sternen und/oder Rezensionstexten helfen den Verbrauchern dabei, sich vor der Bestellung bzw. dem Vertragsschluss ein konkretes Bild von dem Produkt oder dem anbietenden Unternehmen zu machen, verschiedene Angebote miteinander zu vergleichen, offene Fragen zu klären und schließlich eine Entscheidung zu treffen. Nutzerbewertungen im Internet übernehmen damit in der digitalen Welt häufig die Funktion der klassischen Entscheidungshilfen in der analogen Welt wie der eigenen Anschauung oder der persönlichen Beratung im Geschäft. Zahlreiche Verbraucherbefragungen und Studien zeigen, dass Nutzerbewertungen von der Mehrheit der Verbraucher im Internet gelesen werden und neben dem Preis das wichtigste Entscheidungskriterium darstellen.¹

Wenn aber die angezeigten Nutzerbewertungen nicht in der von den Verbrauchern erwarteten Weise entstanden sind, sondern z. B. gar nicht auf einer echten Nutzererfahrung beruhen oder manipuliert wurden, werden die Verbraucher über die Qualität oder die Eigenschaften der bewerteten Produkte und Unternehmen (zusammengefasst „Items“) getäuscht und treffen daraufhin möglicherweise eine falsche Entscheidung. Eine solche Irreführung stellt einen Verstoß gegen die verbraucherrechtlichen Vorschriften des Gesetzes gegen den unlauteren Wettbewerb (UWG)² dar. Während betroffene Portale und Anbieter aufgrund ihrer Informationen und Möglichkeiten teilweise selbst gegen nicht-authentische oder manipulierte Nutzerbewertungen vorgehen können, ist der Verbraucher insbesondere gegenüber systematischen Verzerrungen aufgrund von Manipulationen oder mangelhaften Kontrollen der Nutzerbewertungen nahezu machtlos.

Das Problem ist bekannt. Bei den Verbraucherzentralen der Bundesländer gehen regelmäßig Beschwerden zu verschiedenen Aspekten rund um das Thema ein. Zahlreiche Studien, Verfahren und Berichte beschäftigen sich in den letzten Jahren allein in Deutschland mit verschiedenen Aspekten von „Fake-Bewertungen“ oder „gefälschten Bewertungen“. Der Marktwächter Digitale Welt beim vzbv veröffentlichte im April 2018 eine Studie, wie wirksam verschiedene Online-Portale „gefälschte“ Bewertungen bekämpfen.³ Die Wettbewerbskommission Baden-Württemberg forderte in einer Stellungnahme im Januar 2020, gekaufte Bewertungen im Internet zu verbieten.⁴ Der Sachverständigenrat für Verbraucherfragen hat sich mit der Behandlung von Nutzerbewertungen im Lauterkeitsrecht als Teil einer breit angelegten Stellungnahme zur Plattformregulierung

¹ Siehe z. B. <https://www.bitkom-research.de/de/pressemitteilung/jeder-zweite-liest-online-bewertungen-vor-dem-kauf> (abgerufen am 01.10.2020).

² Gesetz gegen den unlauteren Wettbewerb in der Fassung der Bek. v. 03.03.2010 (BGBl. I S. 254), zuletzt geändert durch Art. 5 des Gesetzes v. 18.04.2019 (BGBl. I S. 466) – UWG.

³ vzbv, Fälschungen bei Bewertungen – Bekämpfen Online-Portale sie wirksam?, April 2018.

⁴ *Verbraucherkommission Baden-Württemberg*, Stellungnahme Nr. 51/2020: Gekaufte Bewertungen im Internet verbieten, 13.01.2020.

auseinander gesetzt.⁵ Eine große Zahl zivilrechtlicher Verfahren, teilweise bis zum Bundesgerichtshof, hat sich mit möglichen Rechtsverstößen im Zusammenhang mit Nutzerbewertungen befasst. Diese Verfahren wurden häufig von den betroffenen Unternehmen gegen die die Bewertungen darstellenden Portale angestrengt.⁶ Teilweise gehen zudem Portalbetreiber gegen Unternehmen vor, die gewerbsmäßig Bewertungen verkaufen bzw. vermitteln.⁷ Das Problem der „Fake-Bewertungen“ spielt inzwischen aber auch in der öffentlichen Diskussion eine große Rolle und ist immer wieder Thema in (kritischen) Medienberichten.⁸

Im Ausland existieren ebenfalls zahlreiche Verfahren und Untersuchungen zum Thema Nutzerbewertungen: Bereits 2011 verhängte die australische Wettbewerbs- und Verbraucherschutzbehörde ACCC ein Bußgeld wegen gefälschter Bewertungen.⁹ Die britische CMA hat während ihrer einjährigen Präsidentschaft im ICPEN¹⁰ 2015/16 das Thema „Online Reviews“ als Schwerpunkt gesetzt, woraufhin in vielen Ländern bis heute Aktivitäten zu dem Thema durchgeführt werden, u. a. von der CMA selbst und der amerikanischen FTC.¹¹

Angesichts der großen Bedeutung von Nutzerbewertungen für den Verbraucher und der gleichzeitig bestehenden Probleme im Hinblick auf „Fake-Bewertungen“ gibt es zudem auf politischer Ebene verschiedene Ansätze, das Thema zu erfassen: So hat im Mai 2017 die Weltorganisation des Tourismus der Vereinten Nationen (UNWTO) bereits „Recommendations on the Responsible Use of Ratings and Reviews on Digital Platforms“ veröffentlicht.¹² Im März 2019 wurde die deutsche Fassung der Norm DIN ISO 20488 „Kundenbewertungen – Grundsätze und Anforderungen für die Erhebung, Moderation und Veröffentlichung“ veröffentlicht,

⁵ *Specht-Riemenschneider*, Stellungnahme des Sachverständigenrats für Verbraucherfragen: Grundlegung einer verbrauchergerichten Regulierung interaktionsmitteler Plattformfunktionalitäten, S. 43 ff., Juni 2020.

⁶ Vgl. z. B. BGH, Urt. v. 14.01.2020, Az. VI ZR 496/18 (Yelp); BGH, Urt. v. 20.02.2020, Az. I ZR 193/18 - *Kundenbewertungen auf Amazon*.

⁷ HolidayCheck, Pressemitteilung v. 13.03.2019, *Im Namen der Urlauber: HolidayCheck sagt Bewertungsbetrügern den Kampf an*; Handelsblatt, *Amazon geht gegen gekaufte Produktbewertungen vor*, 05.03.2019.

⁸ Vgl. zum Beispiel: Frankfurter Allgemeine Sonntagszeitung, *Schummelei mit Sternen*, 06.10.2019; Der Spiegel, *Fünf Sterne für nix*, 12.10.2019; Wirtschaftswoche, *Die Bewertungsmafia*, 17.10.2019; Süddeutsche Zeitung, *Zehnmal Lob für 118,88 Euro*, 16.12.2019; Die Zeit, *Tausche Sterne gegen Ware*, 16.04.2020.

⁹ <https://www.accc.gov.au/media-release/accc-removalist-admits-publishing-false-testimonials> (abgerufen am 01.10.2020).

¹⁰ Das International Consumer Protection and Enforcement Network (ICPEN) ist ein globales Netzwerk von Verbraucherschutzbehörden, das vor allem die internationale Zusammenarbeit der Verbraucherschutz- und Rechtsdurchsetzungsbehörden fördert.

¹¹ <https://www.gov.uk/government/news/cma-investigates-misleading-online-reviews> (abgerufen am 01.10.2020); <https://www.gov.uk/government/news/cma-takes-enforcement-action-against-fake-online-reviews>; <https://www.gov.uk/government/news/facebook-and-ebay-pledge-to-combat-trading-in-fake-reviews>; <https://www.ftc.gov/news-events/press-releases/2019/10/devumi-owner-ceo-settle-ftc-charges-they-sold-fake-indicators>; <https://www.ftc.gov/news-events/press-releases/2019/02/ftc-brings-first-case-challenging-fake-paid-reviews-independent> (alle abgerufen am 01.10.2020).

¹² Siehe <https://www.unwto.org/archive/europe/press-release/2017-05-01/unwto-international-congress-ethics-presents-recommendations-responsible-us> (abgerufen am 01.10.2020).

die insbesondere Empfehlungen zur Ausgestaltung von Bewertungsprozessen enthält und die unter Mitwirkung verschiedener Portale entstanden ist.¹³ Auf legislativer Ebene wurden zuletzt im Rahmen des EU-Projekts „New Deal for Consumers“ im November 2019 einige Vorschriften zu Nutzerbewertungen in die UGP-Richtlinie¹⁴ aufgenommen.¹⁵ Auch in den Überlegungen der Kommission zur Verantwortung von Plattformen im Rahmen des Vorschlags für einen „Digital Services Act“ spielen Nutzerbewertungen eine Rolle.

Vor dem Hintergrund der Bedeutung, der Aktualität, aber auch der Komplexität dieser Thematik hat das Bundeskartellamt im Mai 2019 die Sektoruntersuchung Nutzerbewertungen eingeleitet. Durch die unmittelbare und ausführliche Befragung zahlreicher Marktteilnehmer, die Auswertung von Studien und Rechtsprechungspraxis sowie eine differenzierte Betrachtung der bestehenden Problemkonstellationen soll die Sektoruntersuchung dazu beitragen, Verbraucher, Unternehmen und Politik besser über die Hintergründe und die Zusammenhänge in Bezug auf das Thema Nutzerbewertungen zu informieren, mutmaßliche Verbraucherrechtsprobleme aufzuklären und zu konkretisieren und nicht zuletzt mittels konkreter Lösungsansätze die Durchsetzung des Verbraucherschutzes in diesem Bereich zu verbessern.

B.2. Gegenstand der Sektoruntersuchung Nutzerbewertungen

Der hier zu untersuchende Wirtschaftszweig im Sinne von § 32e Abs. 5 GWB¹⁶ umfasst Dienstleistungen, die im Zusammenhang mit der Erstellung, der Organisation und der Darstellung von Nutzerbewertungen auf Online-Portalen erbracht werden, wenn diese Bewertungen dazu geeignet sind, geschäftliche Entscheidungen der Verbraucher zu beeinflussen, sei es auf dem Portal selbst oder an anderer Stelle. Die Begriffe „Nutzerbewertungen“ oder „Bewertungen“ werden dabei als Oberbegriffe für die vergebenen Sterne, Noten etc. sowie die verfassten Rezensionstexte verwendet. Als geschäftliche Entscheidung wird im Rahmen dieser Sektoruntersuchung die entgeltliche Bestellung oder der Vertragsschluss bezüglich einer Ware, einer Dienstleistung oder eines Werkvertrags angesehen.

In dem untersuchten Wirtschaftszweig sind in erster Linie Portale tätig, die Nutzerbewertungen anzeigen und häufig ein eigenes Nutzerbewertungssystem betreiben oder dieses von dritten Unternehmen betreiben las-

¹³ Zu beziehen über: <https://www.beuth.de/de/norm/din-iso-20488/300183899> (abgerufen am 01.10.2020).

¹⁴ Richtlinie 2005/29/EG des Europäischen Parlaments und des Rates vom 11.05.2005 über unlautere Geschäftspraktiken im binnenmarktinternen Geschäftsverkehr zwischen Unternehmen und Verbrauchern und zur Änderung der Richtlinie 84/450/EWG des Rates, der Richtlinien 97/7/EG, 98/27/EG und 2002/65/EG des Europäischen Parlaments und des Rates sowie der Verordnung (EG) Nr. 2006/2004 des Europäischen Parlaments und des Rates (Richtlinie über unlautere Geschäftspraktiken), Abl. EU Nr. L 149 v. 11.06.2005, S. 22 (kurz UGP-Richtlinie).

¹⁵ Siehe Art. 3 der Richtlinie (EU) 2019/2161 des Europäischen Parlaments und des Rates v. 27.11.2019 zur Änderung der Richtlinie 93/13/EWG des Rates und der Richtlinien 98/6/EG, 2005/29/EG und 2011/83/EU des Europäischen Parlaments und des Rates zur besseren Durchsetzung und Modernisierung der Verbraucherschutzvorschriften der Union, Abl. EU Nr. L 328 v. 18.12.2019, S. 7.

¹⁶ Gesetz gegen Wettbewerbsbeschränkungen in der Fassung der Bek. v. 26.06.2013 (BGBl. I S. 1750, 3245), zuletzt geändert durch Art. 1 des Gesetzes v. 25.05.2020 (BGBl. I S. 1067) - GWB.

sen. Bei diesen Portalen kann es sich um Buchungsportale, Plattformen, Vergleichsportale, Bewertungsportale, Online-Shops oder andere Portale handeln. Daneben gehören zum untersuchten Wirtschaftszweig solche Unternehmen, die Dienstleistungen rund um Nutzerbewertungen anbieten, sowie Unternehmen, die Nutzerbewertungen gegen Entgelt verkaufen oder vermitteln. Um ein vollständiges Bild des Wirtschaftszweigs zu erhalten, ist schließlich auch die Berücksichtigung der Interessenlagen von Anbietern und Verbrauchern erforderlich, wobei diese beiden Gruppen im Rahmen der Sektoruntersuchung nicht explizit befragt wurden.

Die ausführliche Befragung der Portale wurde im Rahmen der Sektoruntersuchung auf diejenigen Bereiche fokussiert, in denen Nutzerbewertungen einen besonders großen Einfluss auf die Entscheidung des Verbrauchers haben. Dies betraf die Bereiche Ärzte, Anwälte/Finanzberater, Apps, Autohändler, Baumarkt, Elektronik, Filme, Generalisten, Handwerker, Kleidung, Medikamente, Reisen/Urlaub, Restaurants, Spiele, Supplements und Veranstaltungen (s. Kapitel C.1.1.).

Die inhaltliche Analyse der Sektoruntersuchung umfasst im Wesentlichen drei Themenfelder: Zunächst soll untersucht werden, wie die von den Portalen betriebenen Bewertungssysteme funktionieren und inwieweit deren Ausgestaltung geeignet ist, die Zahl „gefälschter“ Nutzerbewertungen möglichst systematisch gering zu halten. Des Weiteren soll das Spannungsfeld der unterschiedlichen Interessenlagen verschiedener Marktteilnehmer, d. h. der Portale, der Anbieter, der Verbraucher aber auch der Bewertungsvermittler in Bezug auf Nutzerbewertungen näher beleuchtet und analysiert werden. Und schließlich zielt die Sektoruntersuchung vor diesem Hintergrund darauf ab, eine differenzierte Betrachtung des häufig pauschal verurteilten Problems „Fake-Bewertungen“ vorzunehmen und basierend darauf Lösungsansätze zu formulieren.

Die Darstellung der Ermittlungs- und Analyse-Ergebnisse gibt jeweils den Stand zum Zeitpunkt der Befragung bzw. der Auswertung wieder. Angesichts der ständigen Veränderung und Entwicklung der Marktverhältnisse in der digitalen Welt erheben die Ausführungen keinen Anspruch auf vollständige Aktualität.

C. Ermittlungen

Nach dem GWB kann das Bundeskartellamt im Rahmen einer verbraucherrechtlichen Sektoruntersuchung mit Ausnahme von Nachprüfungen, Durchsuchungen und Beschlagnahmen alle Ermittlungen führen und alle Beweise erheben, die erforderlich sind (§ 57 Abs. 1, § 32e Abs. 4, Abs. 5 S. 3 GWB). Insbesondere hat das Amt die Möglichkeit, zur Erfüllung seiner Aufgaben von Unternehmen und Vereinigungen von Unternehmen mit Sitz in Deutschland Auskünfte über ihre wirtschaftlichen Verhältnisse und die Herausgabe von Unterlagen zu verlangen (§ 59 Abs. 1, § 32e Abs. 4 GWB). Die Unternehmen sind in diesem Fall unter Bußgeldandrohung zur wahrheitsgemäßen Auskunftserteilung verpflichtet und können die Auskunft auch nicht unter Hinweis auf Betriebs- und Geschäftsgeheimnisse verweigern (§ 59 Abs. 2 i. V. m. § 80 Abs. 2 Nr. 6 GWB).

Die im Rahmen der Sektoruntersuchung seitens des Bundeskartellamtes durchgeführten Ermittlungen zielten nicht darauf ab, endgültig gerichtsfest festzustellen, ob ein bestimmtes Verhalten eines Unternehmens einen Verstoß gegen verbraucherrechtliche Vorschriften und dabei insbesondere gegen UWG-Bestimmungen darstellt. Denn für die Feststellung eines konkreten Rechtsverstößes im Einzelfall wären der individuelle, unternehmensspezifische Sachverhalt sowie ggfs. die Erwartungen, die der Verbraucher an die dort dargestellten Bewertungen hat, zu ermitteln. Dies würde aber die gegenwärtigen Befugnisse des Bundeskartellamtes im Verbraucherrecht überdehnen, die nach dem Willen des Gesetzgebers konkrete Eingriffsbefugnisse nicht umfassen. Für die Zwecke der Sektoruntersuchung Nutzerbewertungen waren zudem Ermittlungen zu Verbrauchererwartungen (bspw. durch eine direkte Befragung der Verbraucher) nicht erforderlich, sodass sich das Bundeskartellamt vorliegend auf die Befragung der Portale sowie weiterer Marktteilnehmer konzentriert hat.

Die Ermittlungen bestanden vor allem in der unmittelbaren Befragung zahlreicher reichweitenstarker Portale, die Nutzerbewertungen auf ihrer Website anzeigen, die für die wirtschaftlichen Entscheidungen der Verbraucher von Bedeutung sind. Ergänzend dazu hat das Bundeskartellamt mehrere Vermittler von Nutzerbewertungen schriftlich um Stellungnahme gebeten. Schließlich wurden während der gesamten Sektoruntersuchung Gespräche mit verschiedenen Marktteilnehmern sowie Verbänden und Behörden geführt, die ebenfalls mit dem Thema Nutzerbewertungen befasst sind. Kurz vor Veröffentlichung der ersten Erkenntnisse im Konsultationspapier haben die darin namentlich erwähnten Unternehmen die Gelegenheit erhalten, sie betreffende Textausschnitte zu sichten und zu kommentieren, was im Text entsprechend berücksichtigt wurde. Neben diesen unmittelbaren Ermittlungen hat das Bundeskartellamt im Rahmen der Sektoruntersuchung auch eine Vielzahl von in- und ausländischen Studien sowie Urteilen verschiedener deutscher Gerichte zum Themenkomplex Nutzerbewertungen ausgewertet und die daraus gewonnenen Erkenntnisse in den entsprechenden Kapiteln eingearbeitet.

C.1. Befragung von Portalen, die Nutzerbewertungen anzeigen

Auf Basis der für die Festlegung des Untersuchungsgegenstands vorgenommenen Abgrenzung (s. Kapitel B.2.) hat das Bundeskartellamt zwischen Oktober und Dezember 2019 eine breit angelegte schriftliche Befragung von Portalen durchgeführt, die auf ihrer Webseite selbst Nutzerbewertungen anzeigen.

C.1.1. Auswahl der zu befragenden Portale

Der Schwerpunkt der Ermittlungen sollte dabei auf der Befragung solcher Portale liegen, deren angezeigte Nutzerbewertungen für deutsche Verbraucher und deren wirtschaftliche Entscheidungen von besonderer Bedeutung sind. Die Zahl der zu befragenden Portale sollte dabei einerseits hinreichend groß sein, um für den Wirtschaftszweig aussagekräftige Ergebnisse zu erhalten, andererseits jedoch innerhalb eines überschaubaren Zeitrahmens für die Beschlussabteilung Wettbewerbs- und Verbraucherschutz des Bundeskartellamts handhabbar sein.

Um einen Überblick über die reichweitenstärksten Portale mit Nutzerbewertungen in Deutschland zu erhalten, hat das Bundeskartellamt in einem ersten Schritt die Übersicht der Top 500 Websites auf Alexa.com¹⁷ betrachtet und daraus diejenigen Websites ausgewählt, die Nutzerbewertungen anzeigen und bei denen diese Bewertungen auch Einfluss auf geschäftliche Handlungen der Verbraucher haben können. Die so ausgewählten Portale konnten insgesamt zwölf verschiedenen Branchen zugeordnet werden: Ärzte, Autohändler, Baummarktartikel, Elektronikartikel, Filme, Generalisten (breites Sortiment), Kleidung, Medikamente, Reisen, Restaurants, Spiele und Veranstaltungen. Aufgrund der anzunehmenden größeren praktischen Bedeutung für Verbraucher wurde diese Liste um vier weitere Branchenkategorien ergänzt: Anwälte/Finanzberater, Apps, Handwerker und Supplements (Nahrungsergänzungsmittel). Bei den damit insgesamt festgelegten 16 relevanten Branchen steht teilweise die Bewertung von Produkten (z. B. Baummarktartikel), teilweise die Bewertung von Unternehmen (z. B. Handwerker) und teilweise die Bewertung von Apps (z. B. Wetter-App) im Vordergrund.

Um auch solche Portale zu erfassen, die für den Verbraucher eine große Rolle spielen, die jedoch nicht zu den Top 500 Websites in Deutschland zählen, hat das Bundeskartellamt in einem zweiten Schritt für jede der 16 Branchen auf den Suchmaschinen Google, Bing und Ecosia jeweils eine Recherche mit drei Suchbegriffen durchgeführt (z. B. bei Medikamenten: „Bewertung Medikamente“, „Medikamente“ und „Medikamente bestellen“). Die ersten beiden Seiten der Suchergebnisse wurden jeweils auf Websites mit Nutzerbewertungen durchsucht.¹⁸

In einem dritten Schritt hat das Bundeskartellamt schließlich für die auf diese Weise zusätzlich erfassten Portale jeweils die Zahl der dort angezeigten Bewertungen für ein bestimmtes, typisches Produkt bzw. Unternehmen ermittelt sowie die Zahl der Facebook-Likes für das jeweilige Portal betrachtet.¹⁹ Für jede der 16 Branchen wurden daraufhin die beiden Portale mit den meisten Bewertungen bzw. Likes ausgewählt. Auf diese Weise konnten von den zusätzlich erfassten Portalen jeweils die für den Verbraucher relevantesten ausgewählt und gleichzeitig die Gesamtzahl der zu befragenden Portale in einem handhabbaren Rahmen gehalten werden. Das Portal Yelp.de hat keines der vorgenannten Auswahlkriterien erfüllt, wurde jedoch

¹⁷ Alexa ist ein Online-Dienst, der Daten über Seitenabrufe von Websites sammelt und darstellt. Mittels verschiedener Faktoren wird dort ein sog. „Alexa Rank“ errechnet. Besonders entscheidend ist dabei die festgestellte Besucheranzahl. Datum der Abfrage: 03.06.2019.

¹⁸ Zeitraum der Suchabfragen: 18.06.2019 - 21.06.2019; Standort der Suchabfrage: Bonn.

¹⁹ Zeitraum der Suchabfragen: 24.06.2019 – 18.07.2019; Standort der Suchabfrage: Bonn.

aufgrund seiner zuvor großen Relevanz in der öffentlichen Diskussion²⁰ sowie seiner Beteiligung an mehreren Gerichtsverfahren²¹ zum Thema Nutzerbewertungen ebenfalls befragt.

Aufgrund der beschriebenen Vorgehensweise hat das Bundeskartellamt insgesamt 66 Portale für die Befragung ausgewählt (siehe Übersicht „Befragte Portale“ im Anhang).

C.1.2. Durchführung der Befragung

Bei der Entwicklung der Fragebögen für die Portale hat das Bundeskartellamt alle wesentlichen, in bisherigen Studien, Gerichtsverfahren und Beschwerden thematisierten Problemfelder im Bereich Nutzerbewertungen berücksichtigt, sofern diese für den Verbraucher relevant sind. Die Strukturierung der Fragebögen orientierte sich an den drei Stufen, die eine Nutzerbewertung typischerweise durchläuft - Erfassung, Filterung, Veröffentlichung - und die vom betreffenden Portal grundsätzlich beeinflussbar sind.²²

Angesichts der Unterschiede, die bei der Bewertung von Produkten, von Unternehmen bzw. von Apps bestehen, hat das Bundeskartellamt für jeden dieser drei Bereiche einen separaten Fragebogen erstellt. Je nach Art bzw. Schwerpunkt der angezeigten Nutzerbewertungen hatten die ausgewählten Portale dann entweder den Fragebogen für Produktbewertungen (Elektronik, Filme, Kleidung etc.), für Unternehmensbewertungen (Händler, Ärzte, Hotels etc.) oder für App-Bewertungen auszufüllen; von einzelnen Portalen waren zwei verschiedene Fragebögen zu beantworten.

Im Oktober 2019 hat das Bundeskartellamt die 66 ausgewählten Portale schriftlich um Auskunft gebeten und den bzw. die Fragebögen mittels eines Links bereitgestellt.²³ Sofern dem Bundeskartellamt eine deutsche Postanschrift vorlag, wurde den betreffenden Portalen ein Auskunftsbeschluss zugestellt, sodass sie zur Beantwortung des Fragebogens verpflichtet waren. Für insgesamt acht Portale wurde das Schreiben mit dem Link zum Fragebogen an deren ausländische Postanschrift geschickt und der Auskunftsbeschluss lediglich informationshalber übersandt. Da die betreffenden Unternehmen nicht über eine deutsche Niederlassung verfügten, konnte das Bundeskartellamt diese nach geltender Rechtslage nicht zur Auskunft verpflichten.

Ein Portal (Pearl) hat gegen den zugestellten Auskunftsbeschluss Beschwerde eingelegt, diese jedoch kurz vor der vom Oberlandesgericht Düsseldorf angesetzten mündlichen Verhandlung im März 2020 zurückgenommen und schließlich im Mai 2020 den Fragebogen beantwortet. In seinem Kostenbeschluss hat das Oberlandesgericht Düsseldorf die formelle und materielle Rechtmäßigkeit des Auskunftsbeschlusses bestätigt.²⁴

²⁰ Vgl. z. B. Süddeutsche Zeitung, *Die Macht der Sterne*, 30.07.2019.

²¹ Vgl. u. a. Entscheidungen KG Berlin, Urt. v. 10.12.2015, Az. 10 U 26/15; Hanseatisches Oberlandesgericht Hamburg, Urt. v. 10.11.2015, Az. 7 U 18/15; BGH,-Urt. v. 14.01.2020, Az. VI ZR 496/18 (u.a.).

²² Dies entspricht auch der Struktur der oben bereits erwähnten Norm DIN-ISO-20488.

²³ Bundeskartellamt, Sektoruntersuchung Nutzerbewertungen, <https://www.bundeskartellamt.de/nutzerbewertungen.html> (abgerufen am 01.10.2020).

²⁴ OLG Düsseldorf, Beschl. v. 18.03.2020, Az. VI-Kart 7/19 (V) (nicht veröffentlicht).

Einem Portal (Imdb.com) hat das Bundeskartellamt die Beantwortung des Fragebogens erlassen, nachdem das Portal darauf aufmerksam gemacht hatte, dass seine Website ganz überwiegend englischsprachige Texte beinhaltet. Von den auf freiwilliger Basis im Ausland befragten acht Portalen haben drei den Fragebogen nicht beantwortet (de.Gearbest.com, Myprotein.de und de.Shein.com). Insgesamt hat das Bundeskartellamt damit von 62 Portalen Antworten erhalten, darunter fünf aus dem Ausland. Zum Fragebogen Produktbewertungen sind 28 Antworten eingegangen, zum Fragebogen Unternehmensbewertungen 36 Antworten und zum Fragebogen App-Bewertungen 3 Antworten. Ergebnis dieser Ermittlungen waren folglich 67 beantwortete Fragebögen (siehe Übersicht „Befragte Portale“ im Anhang). Die befragten Portale haben zudem dem Fragebogen entsprechend zahlreiche Verträge, interne Richtlinien sowie Gerichtsurteile zum Thema Nutzerbewertungen übersandt. Angesichts einer Rücklaufquote (inkl. Ausland) von über 90 % sowie der Tatsache, dass für alle drei Fragebögen und alle betrachteten Branchen jeweils die relevantesten Portale geantwortet und aussagekräftige Unterlagen vorgelegt haben, konnte die Auswertung insgesamt auf einer sehr soliden Datenbasis erfolgen.

C.2. Stellungnahmen von Bewertungsvermittlern

Neben der überwiegend per Auskunftsbefragung erfolgten Befragung von Portalen hat sich das Bundeskartellamt auch an solche Unternehmen gewandt, die im Internet Nutzerbewertungen zum Kauf anbieten bzw. die eine entgeltliche Vermittlung von (privaten) „Produkttestern“ vornehmen. Derartige Angebote richten sich zumeist an Händler, Hersteller oder Dienstleister, die die Darstellung und den Verkauf ihrer eigenen Leistungen über Plattformen im Internet verbessern wollen. Die Verfasser der Bewertungen erhalten hierfür i. d. R. eine Vergütung. Die Vermittlung von derartig incentivierten Nutzerbewertungen kann jedoch mit verschiedenen rechtlichen Problemen verbunden sein und wurde zuletzt häufiger öffentlich kritisiert bzw. gerichtlich verfolgt. Die betreffenden Bewertungsvermittler sind teilweise auch nur schwer erreichbar. Viele verfügen trotz Website in deutscher Sprache nur über einen Unternehmenssitz im Ausland. Teilweise erfolgt die Vermittlung auch gar nicht offiziell über Unternehmen, sondern über Gruppen in (sozialen) Netzwerken, deren tatsächliche Organisatoren i. d. R. nicht erkennbar sind.

Auf der Basis von Eingaben durch Verbraucher und Unternehmen sowie eigener Internet-Recherchen hat das Bundeskartellamt im November 2019 insgesamt 30 Betreiber von verschiedenen Websites angeschrieben, die gewerbsmäßig die Vermittlung von Nutzerbewertungen für deutsche Plattformen (u. a. Amazon, Google, Ebay, Facebook)²⁵ anbieten. Die Bewertungsvermittler wurden über Inhalt und Ziele der Sektoruntersuchung informiert und erhielten Gelegenheit, zu verschiedenen konkreten Fragen sowie zur öffentlichen Kritik an ihrem Geschäftsmodell allgemein Stellung zu nehmen (siehe Übersicht „Befragte Bewertungsvermittler“ im

²⁵ Zur besseren Lesbarkeit werden im Folgenden Portale mit bekannten oder eindeutigen Namen (z. B. „Amazon“, „Google Search“ oder „Facebook“) teilweise ohne die vollständige URL genannt. Die Nennung der vollständigen URL bei anderen Portalen (z. B. „Anwalt.de“ oder „Mobile.de“) erfolgt lediglich zur Vermeidung von Missverständnissen.

Anhang). Die Reaktionen auf diese Anschreiben fielen sehr unterschiedlich aus. In sieben Fällen kam das Anschreiben mit dem Vermerk „unzustellbar“ zurück, obwohl die gewählte Anschrift der auf der Homepage des Bewertungsvermittlers genannten Postadresse entsprach. Bis Dezember 2019 hat das Bundeskartellamt zu dem Schreiben von insgesamt sieben Unternehmen inhaltliche Rückmeldungen erhalten, entweder schriftlich, per E-Mail, telefonisch oder in persönlichen Gesprächen. Im Rahmen der oben erläuterten Befragung von Portalen, die Nutzerbewertungen anzeigen, hatten einige dieser Portale ihrerseits ebenfalls die ihnen bekannten Bewertungsvermittler benannt. Hierbei handelte es sich ganz überwiegend um bereits berücksichtigte Unternehmen oder um nicht-deutschsprachige Websites.

Die Rückmeldungen der Bewertungsvermittler lieferten wertvolle Erkenntnisse zu den unterschiedlichen Systemen sowie den aktuellen Herausforderungen auf Seiten dieser Unternehmen. Zur Erfassung der Rolle von Produkttestergruppen in (sozialen) Netzwerken hat das Bundeskartellamt ergänzend dazu eigene Internet-Recherchen durchgeführt.

C.3. Gespräche mit Marktteilnehmern und weiteren Institutionen

In allen Phasen der Sektoruntersuchung hat das Bundeskartellamt zahlreiche Gespräche mit unterschiedlichen Marktteilnehmern und weiteren Institutionen geführt. Dazu gehörten vor allem Portale und andere (z. B. als Dienstleister für Nutzerbewertungen tätige) Unternehmen, Verbände sowie ausländische Wettbewerbs- und Verbraucherschutzbehörden, die ebenfalls Untersuchungen oder Verfahren im Bereich Nutzerbewertungen durchführen. Die Gespräche fanden telefonisch, per Videokonferenz oder vor Ort im Bundeskartellamt statt und brachten zahlreiche zusätzliche Erkenntnisse und Hintergrundinformationen über die Funktionsweise von Bewertungssystemen, die grundsätzliche Bedeutung von Nutzerbewertungen sowie die damit verbundenen Probleme und Rechtsfragen. Zu den Marktteilnehmern, mit denen das Bundeskartellamt Gespräche geführt hat, gehörten u. a. Amazon, Bazaarvoice, Ebay, eKomi, Google, GoLocal, Review Meta, Otto und Yelp. Zu den Gesprächspartnern auf Seiten der Verbände und Behörden gehörten Bitkom, der Händlerbund, der Verbraucherzentrale Bundesverband (vzbv), die irische Competition and Consumer Protection Commission (CCPC), die britische Consumer and Markets Authority (CMA) und die US-amerikanische Federal Trade Commission (FTC).

C.4. Konsultation

Mit der Veröffentlichung des Konsultationspapiers zur vorliegenden Sektoruntersuchung Nutzerbewertungen im Juni 2020 hat das Bundeskartellamt den befragten Unternehmen sowie weiteren interessierten Kreisen die Gelegenheit gegeben, zu den im Konsultationspapier vorgestellten verbraucherrechtlichen Einschätzungen und Lösungsansätzen Stellung zu nehmen.²⁶ Insgesamt sind elf schriftliche Stellungnahmen von ver-

²⁶ Siehe Pressemitteilung vom 18.06.2020, hier abrufbar: https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2020/18_06_2020_SU_Nutzerbewertungen_Konsultation.html.

schiedenen Marktteilnehmern und Institutionen eingegangen. Dazu gehörten neben betroffenen Unternehmen der vzbv, das Ministerium für ländlichen Raum und Verbraucherschutz Baden-Württemberg sowie die Verbraucherkommission Baden-Württemberg. Sowohl hinsichtlich des Umfangs als auch des Inhalts wiesen die eingegangenen Stellungnahmen naturgemäß große Unterschiede auf. Das Bundeskartellamt hat eine Würdigung aller Stellungnahmen vorgenommen und diese soweit möglich in den Bericht aufgenommen, dabei aber mit Ausnahme von Stellungnahmen durch öffentliche Institutionen weitgehend auf eine Namensnennung verzichtet.

Sowohl das Ministerium für ländlichen Raum und Verbraucherschutz als auch die Verbraucherkommission Baden-Württemberg betonen die Bedeutung des Themas. Das Ministerium regt an, die Untersuchung auf Dienste wie Instagram, Pinterest und Youtube auszuweiten und außerdem eine Verbraucherbefragung durchzuführen, um Verbrauchererwartungen und tatsächliche Situation im Bereich Nutzerbewertungen gegenüberstellen zu können. Beide Stellungnahmen regen eine Erweiterung der Durchsetzungsbefugnisse des Bundeskartellamts im Bereich des Verbraucherschutzes an, um die zivile Rechtsdurchsetzung zu ergänzen.

D. Funktionsweisen von Bewertungssystemen

Der Umgang der Portale mit ihren Nutzerbewertungen bildet in der vorliegenden Sektoruntersuchung den Schwerpunkt der Ermittlungen. Untersucht wurden Bewertungen für Produkte, Unternehmen und Apps. Im Fokus stand dabei nicht die einzelne Bewertung, sondern die Gesamtheit aller Regelungen und Maßnahmen eines Portals im Umgang mit Bewertungen, weshalb im Folgenden der Begriff des Bewertungssystems verwendet wird.

Ein solches Bewertungssystem besteht grundsätzlich aus drei Komponenten, zu denen sich jeweils verschiedene Fragestellungen ergeben, die das Bundeskartellamt im Rahmen seiner Ermittlungen bei den ausgewählten 66 Portalen untersucht hat. Die erste Komponente ist die **Erfassung** von Bewertungen: Welche Leistungen können auf dem Portal bewertet werden? Wer kann unter welchen Voraussetzungen Bewertungen abgeben? Welche Maßnahmen ergreifen die Portale, um Bewertungen zu erhalten? Die zweite Komponente betrifft die **Filterung** der erfassten Bewertungen: Werden alle eingegangenen Bewertungen manuell geprüft, oder gibt es automatisierte Filter, die eine Vorsortierung vornehmen? Welche Kriterien liegen diesen Filtern jeweils zugrunde? Können auffällige Bewertungen von Nutzern bzw. Anbietern auf dem Portal gemeldet werden? Wie führt das Portal ggfs. die manuelle Kontrolle oder Bearbeitung von gefilterten bzw. gemeldeten Bewertungen durch? In welcher Form werden Verstöße gegen Bewertungsrichtlinien ggfs. sanktioniert? Schließlich besteht die dritte Komponente des Bewertungssystems in der **Darstellung** der ausgewählten Nutzerbewertungen: In welcher Form werden die Bewertungen angezeigt? Werden auch Bewertungen, die von anderen Portalen erfasst und gefiltert wurden, angezeigt? Wie erfolgt die Berechnung der Durchschnittsbewertung oder Gesamtnote? Die nachfolgenden Ausführungen in diesem Kapitel stellen die Ermittlungsergebnisse zu den einzelnen Komponenten der Bewertungssysteme sowie den konkreten Fragestellungen dar und beleuchten die in diesem Zusammenhang identifizierten Probleme.

Einige der befragten Portale haben angegeben, ihr Bewertungssystem nicht selbst, sondern mithilfe eines externen Dienstleisters zu betreiben. Aufgrund der dazu übersandten Unterlagen und ergänzender Gespräche mit Marktteilnehmern konnte das Bundeskartellamt die Vorgehensweisen derartiger Dienstleister in Bezug auf die drei genannten Komponenten bei der Auswertung berücksichtigen. Eingesetzt werden diese Dienstleister insbesondere im Bereich der Produktbewertungen: Mehr als ein Drittel aller Online-Shops, die Produktbewertungen anzeigen, haben teilweise die Erfassung, Filterung und Darstellung von Bewertungen ausgelagert. Der dabei am häufigsten genannte Kooperationspartner war das Unternehmen Bazaarvoice, ein amerikanischer Softwareanbieter und Dienstleister im Bereich Consumer Generated Content (CGC), der auch als Marktführer in diesem Bereich bezeichnet wurde. Im Bereich der Unternehmensbewertungen war die Zusammenarbeit mit einem externen Dienstleister weniger stark verbreitet; hier haben die Ermittlungen ergeben, dass einzelne Portale das von GoLocal betriebene Bewertungssystem nutzen.

D.1. Erfassung von Bewertungen

D.1.1. Welche Leistungen können bewertet werden?

Bei der weit überwiegenden Zahl der befragten Portale können alle der dort dargestellten Items, d. h. alle dargestellten Produkte und Unternehmen, von den Nutzern auch bewertet werden. Ausnahmen gelten bei einigen Portalen für bestimmte Produktbereiche. So haben Portale, die Medikamente und Heilmittel anbieten, darauf hingewiesen, dass aufgrund gesetzlicher Vorgaben zur Werbung für verschreibungspflichtige Arzneimittel diese auf ihrer Seite nicht bewertet werden können (s. § 10 Abs. 1 HWG²⁷). Ein Portal hat zudem berichtet, dass bei ihm die Bewertung von Erotikprodukten nicht möglich sei.

Unterschiede bestehen hingegen im Hinblick auf die vorgelagerte Frage, welche Items überhaupt auf einem Portal dargestellt werden bzw. auf welchem Wege Items auf einem Portal eingestellt werden können, sodass Nutzer dann Bewertungen dafür abgeben können.

Bei den typischen Handels- oder Buchungsportalen, über die der Verbraucher eine Transaktion abschließen kann, erfolgt die Einstellung eines Items in der Regel auf vertraglicher Basis durch Portal und Händler/Anbieter gemeinsam und im beiderseitigen Interesse. Dies gilt für Plattformen wie Ebay, Amazon oder Booking, aber naturgemäß auch für Online-Shops, die von den Anbietern selbst betrieben werden.

Auf Portalen, bei denen die Suche, der Vergleich oder die Bewertung von Items im Vordergrund stehen, haben häufig Anbieter direkt die Möglichkeit, für sich selbst ein Profil anzulegen bzw. dies beim Portal zu beantragen. Der Antrag wird dann von den Portalen unterschiedlich streng geprüft, z. B. daraufhin, ob das Unternehmen überhaupt existiert oder ob es über eine gültige Gewerbeanmeldung verfügt. So können beispielsweise Unternehmen, die bei Google Maps angezeigt werden möchten, dies über eine Anmeldung bei Google My Business tun.²⁸ Google schaltet einen Brancheneintrag frei, wenn es die Identität des Antragstellers überprüft hat. Das Portal Anwalt.de bietet Anwälten die Möglichkeit, gegen eine monatliche Gebühr ein Profil zu erstellen, nur dann werden sie auch auf dem Portal dargestellt.

Bei manchen Portalen können neben den Anbietern auch Nutzer ein Profil für ein Unternehmen oder ein anderes Item anlegen, das noch nicht auf dem Portal sichtbar ist und zu dem dann Bewertungen abgegeben werden können: Google (für Google Maps), Holidaycheck, Kennstduenein.de, Tripadvisor, Trustpilot, Whofinance.de und Yelp gaben an, dass bei ihnen diese Möglichkeit bestehe, wobei teilweise vor Veröffentlichung die Richtigkeit der Angaben noch vom Portal überprüft wird. Die betroffenen Anbieter müssen der Anlage des Profils teilweise nicht zustimmen, können jedoch ggfs. den Eintrag „übernehmen“ und dann je nach Portal Ergänzungen wie z. B. Beratungsschwerpunkte, individuelle Angebote, Öffnungszeiten oder Speisekarten vornehmen.

²⁷ Gesetz über die Werbung auf dem Gebiete des Heilwesens in der Fassung der Bek. v. 19.10.1994 (BGBl. I S. 3068), zuletzt geändert durch Art. 6 des Gesetzes v. 18.04.2020 (BGBl. I S. 960) – Heilmittelwerbeengesetz (HWG).

²⁸ <https://support.google.com/business/answer/2911778> (abgerufen am 01.10.2020)

Schließlich greifen einige Portale auch auf Datensätze Dritter zurück, um die Listen der von ihnen dargestellten Items zu ergänzen und zu vervollständigen, bzw. nutzen sogar nur diesen Weg, um Einträge anzulegen. Das Vergleichsportal Check24 nutzt beispielsweise die Daten eines Informationsdienstleisters aus der Energiebranche, um Anbieter in seinen Vergleich aufnehmen zu können, die nicht selbst die Vermittlung über Check24 beantragt haben. Die Listings bei Jameda und bei Werkenntdenbesten.de basieren u. a. auf Datensätzen eines international tätigen Datendienstleisters; Werkenntdenbesten.de nutzt außerdem Verzeichnisdaten von Telekommunikationsanbietern. Einige Portale gaben zudem an, selbst im Internet nach geeigneten Anbietern zu suchen und diese dann in ihre Listen aufzunehmen. Wenn die entsprechenden Anbieter auf dem Portal eingetragen sind, können sie in der Regel auch direkt bewertet werden.

Das zuletzt beschriebene Geschäftsmodell, Unternehmensnamen bzw. -daten in großem Umfang aus bestehenden Verzeichnissen bzw. aus dem Internet zu „kopieren“, diese in eine eigene Liste zu integrieren und damit eine Bewertungsmöglichkeit mit entsprechendem Ranking dafür anzubieten, wurde im Rahmen der Sektoruntersuchung von einigen Marktteilnehmern kritisiert: Insbesondere wenn das betreibende Portal dann zur Verbesserung der Unternehmensdarstellung eigene Servicepakete zum Kauf anbietet, werde über diese Vorgehensweise letztlich erheblicher Druck auf die Unternehmen zum Geschäftsabschluss ausgeübt.

D.1.2. Wer kann Bewertungen abgeben?

Unterschiedlich streng sind die Portale bei der Frage, welche Voraussetzungen ein Nutzer erfüllen muss, um eine Bewertung abgeben zu können. Teilweise ergeben sich diese Unterschiede aus den unterschiedlichen Geschäftsmodellen. So besteht die Möglichkeit, eine Transaktion zur Bedingung für die Abgabe einer Bewertung zu machen, natürlich nur in den Fällen, in denen überhaupt eine Transaktion, also beispielsweise der Kauf eines Produktes, auf einem Portal stattfinden kann. Über den tatsächlichen Besuch eines Restaurants oder einer Arztpraxis ist ein Bewertungsportal dagegen in der Regel nicht automatisch informiert, und ein entsprechender Nachweis wird in der Regel nicht zur Bedingung für die Abgabe einer Bewertung gemacht, denn Portale haben grundsätzlich ein Interesse daran, viele Bewertungen zeigen zu können (s. Kapitel E.4.1.) und versuchen daher, die Hürden für die Abgabe einer Bewertung möglichst niedrig zu halten.

Bei dreizehn der befragten Portale (21 %) in den Bereichen Produkt- und Unternehmensbewertungen können alle Besucher der Website ohne jegliche Voraussetzung eine Bewertung abgeben. Zehn davon sind Handelportale, auf denen Produkte bewertet werden können. Drei sind Portale, auf denen Unternehmen bzw. Dienstleistungen bewertet werden. Alle anderen Portale schränken die Möglichkeit, Bewertungen abzugeben, ein: 46 % der Befragten gaben an, dass Nutzer auf ihrem Portal registriert und angemeldet sein müssen, um eine Bewertung abgeben zu können. Da die reine Anmeldung (z. B. per E-Mail-Adresse) eine vergleichsweise niedrige Hürde ist, werden auch diese Bewertungssysteme zu den **offenen Bewertungssystemen** gezählt.

25 % der Befragten machen den Abschluss einer Transaktion, also z. B. eines Kaufs oder einer Buchung, zur Voraussetzung für die Abgabe einer Bewertung. Auf 13 % der Portale können nur Nutzer, die explizit dazu

aufgefordert wurden, eine Bewertung abgeben. Bei manchen Portalen müssen mehrere der genannten Bedingungen erfüllt sein. In allen diesen Fällen handelt es sich um sog. **geschlossene Bewertungssysteme**.

Neben diesen typischen Bedingungen gelten bei einigen Portalen auch (ergänzende) Sonderregelungen: Auf Amazon können nur solche Kunden bewerten, deren Umsatz bei mindestens 50 € in den letzten 12 Monaten lag.²⁹ Bei Maxdome kann ein Film, der auf einem Smart TV angeschaut wird, erst bewertet werden, wenn er zu mindestens 80 % abgespielt wurde. Acht der befragten Portale (17 %) nutzen eine Software, die die Abgabe einer Bewertung von vornherein verhindert, wenn der damit verbundene Rezensionstext bestimmte Schimpfwörter enthält (vgl. dazu auch Kapitel D.2.).

Auf Appstores können grundsätzlich nur Apps bewertet werden, die heruntergeladen wurden. Auf der Spieleplattform Steam können zudem nur Spiele bewertet werden, die eine bestimmte Zeit lang genutzt wurden. Google wies darauf hin, dass Apps im Playstore nur von Nutzern ab 13 Jahren bewertet werden können.

In der Stellungnahme eines Rechtsanwalts wurde angeregt, ein Mindestalter von 16 Jahren für die Abgabe einer Bewertung gesetzlich festzuschreiben, da diese für den Verfasser rechtlich nachteilhaft sein könne.

D.1.3. Anstupser und Anreize zur Bewertungsabgabe

Grundsätzlich haben Portale ein Interesse daran, möglichst viele Bewertungen zu erhalten, weil sie dadurch mehr Bewertungen anzeigen können, was die Attraktivität ihres Portals für Nutzer erhöht und damit für höhere Umsätze sorgt, auch wenn dieser Effekt bisher vor allem für einzelne Anbieter bzw. Items nachgewiesen wurde (s. Kapitel E.2.1). Durch Anstupser³⁰ sowie verschiedene Anreize zur Bewertungsabgabe kann die Zahl der abgegebenen Bewertungen deutlich erhöhen werden. Dies wurde in der Befragung und in Gesprächen mit Marktteilnehmern mehrfach bestätigt. Die Möglichkeiten der Portale, auf diese Weise mehr Nutzerbewertungen zu generieren, sind allerdings durch rechtliche Vorgaben, insbesondere hinsichtlich Datenschutz und Werbung, eingeschränkt.

Ein einfacher Weg, durch Anstupser mehr Bewertungen zu generieren, ist die explizite Aufforderung an Nutzer, infolge einer Transaktion eine Bewertung abzugeben. Das machen fast alle Portale, die die Möglichkeit dazu haben. Meistens erhalten dabei Nutzer, die dem zugestimmt haben, nach Abschluss einer Transaktion wie z. B. dem Kauf eines Produkts auf einem Portal eine Aufforderung per E-Mail, das Produkt zu bewerten, wobei die Aufforderung je nach Portal bis zu vier Mal erfolgt. Teilweise erhalten Nutzer auch innerhalb der zugehörigen App eine sog. Push Notification. Mehr als die Hälfte der Befragten hat angegeben, darüber hinaus keinerlei Anreize zur Bewertungsabgabe zu setzen.

Auch offline am sog. „touch point“, an dem die Bewerteten und die potentiellen Bewertungsverfasser aufeinander treffen, können solche Anstupser helfen, die Zahl der Bewertungen zu erhöhen. So hat eKomi, ein

²⁹ <https://www.amazon.de/gp/help/customer/display.html?nodeId=201929730> (abgerufen am 01.10.2020).

³⁰ Als Anstupser werden hier Maßnahmen verstanden, die das Verhalten der Nutzer beeinflussen sollen, ohne dass direkte Anreize gesetzt werden (Nudging).

Dienstleister im Bereich Bewertungsmanagement, berichtet, dass das Einrichten solcher Anstupser („Initiativen“) Teil ihrer Angebotspalette sei. Dabei werden z. B. in Arztpraxen am Empfang Karten mit der Bitte um Bewertung, versehen mit einem QR-Code, aufgestellt, um auf diesem Wege mehr Bewertungen zu generieren.

Einige dieser Anstupser haben eine asymmetrische Komponente. So werden häufig Aufforderungen der folgenden Art verwendet: „Sind Sie zufrieden? Schreiben Sie es online. Sind Sie unzufrieden? Schreiben Sie es uns.“ Noch ausgefeilter ist das Vorgehen eines Online-Shops, bei dem alle Kunden per Mail zu einer Sternbewertung aufgefordert, aber nur diejenigen, die positiv bewerten, anschließend um die Abgabe einer ausführlichen Bewertung auf einem Bewertungsportal gebeten werden. Asymmetrisch ist ein Anstupser aber bereits, wenn die Abgabe einer schlechten Bewertung umständlicher ist als die Abgabe einer guten Bewertung: Wenn der Nutzer zufrieden ist, reicht ein Klick, wenn er unzufrieden ist, wird er aufgefordert, dies zu begründen. Diese Methoden führen dazu, dass eher positive als negative Bewertungen abgegeben werden und so ein verzerrtes Bild bezüglich der tatsächlichen Zufriedenheit der Verbraucher entstehen kann.

Neben Anstupsern gewähren einige Portale auch Anreize für eine Bewertung von Items. Dabei geht die Anreizsetzung mit ganz unterschiedlichen Ausgestaltungen einher:

Einen Anreiz, besonders viele, aussagekräftige und mit Bildern versehene Rezensionen zu verfassen, setzen einige der befragten Portale, insbesondere der reinen Bewertungsportale, indem sie eine Bewertungscommunity schaffen. Dabei erhalten Bewertungsverfasser für ihre Rezensionen virtuelle Punkte und bei Erreichen gewisser Punktzahlen virtuelle Auszeichnungen. Die Bewerter können sich mit Hilfe eines aussagekräftigen Namens und eines Profilbildes eine Identität verschaffen, der andere Nutzer wiederum folgen können („Gamification“). Bei Google können Nutzer „local guide“ werden und je nach Anzahl von Beiträgen verschiedene Level erreichen, die durch verschiedenfarbige Sterne am Profilbild gekennzeichnet werden. Die beiden Portale Dasoertliche.de und Dastelefonbuch.de nutzen das Bewertungsmanagement von GoLocal, bei dem ebenfalls Ränge nach der Anzahl an Beiträgen verliehen werden. Der Bewerter startet hier als „Bürger“ und wird dann über 15 Ränge wie „Edler“, „Baron“, „Graf“, „Fürst“ bis zum „Kaiser“ und schließlich zum „Weltbeherrscher“ befördert. Bei Tripadvisor bekommen Teilnehmer am Programm „TripCollective“ Bewerter-Badges über sechs Hierarchiestufen. Yelp bietet Bewertern die Aufnahme in den „Yelp Elite Squad“. Bei Handels- und Verkaufsplattformen wird diese Art des ideellen Anreizes für Bewertungen kaum verwendet. Von den befragten Portalen veröffentlicht nur Amazon ebenfalls ein Ranking seiner Kundenrezensenten („Top-Rezensenten“ und „Hall of Fame-Rezensenten“).³¹ Mehrere Portalbetreiber haben in den Gesprächen mit dem Bundeskartellamt auf die große Bedeutung dieser Communities für das Generieren von Bewertungen hingewiesen.

³¹ Google: https://support.google.com/local-guides/answer/6225846?hl=de&ref_topic=6225845, GoLocal: <https://blog.golocal.de/die-range-auf-golocal/>, Tripadvisor: <https://www.tripadvisor.de/TripCollective>, Yelp: <https://www.yelp.de/elite>, Amazon: <https://www.amazon.de/hz/leaderboard/hall-of-fame> (alle abgerufen am 01.10.2020).

Etwa 20 % aller Befragten verwendet Gutscheine als Anreiz, eine Bewertung abzugeben. Solche finanziellen Anreize gelten als zulässig, solange der Wert des Gutscheins hinreichend gering ist. Ein Portal berichtete, dass es sich bei der Bestimmung der Grenze für einen zulässigen Wert an der Rechtsprechung des Bundesgerichtshofs zu Werbebeigaben eines Apothekers orientiere.³² Bei einem großen Onlinehändler erhalten registrierte Kunden 1 Euro pro Bewertung (maximal 10 Euro pro Jahr). Shopapotheke.de vergibt für Bewertungen mit mindestens 50 Wörtern (Premiumbewertung) einen 5 %-Gutschein für den eigenen Shop. GoLocal vergibt als Anreiz für Neumitglieder für die ersten zehn Bewertungen einen 5-Euro-Amazongutschein, wobei die Vergabe des Gutscheins an bestimmte Qualitätskriterien wie Länge und Aussagekraft geknüpft ist.

Noch beliebter im Bereich der Produktbewertungen sind Gewinnspiele als Anreiz für die Abgabe von Bewertungen, bei Online-Händlern teilweise in Kooperation mit einzelnen Herstellern. Gewinnspiele werden von 40 % der befragten Portale in diesem Bereich zumindest ab und zu eingesetzt. So verlost z. B. Mediamarkt monatlich 100-, 200- und 300-Euro-Coupons unter den Bewertungsverfassern. In diesen Fällen wird die Abgabe einer Bewertung also nicht mit einer unmittelbaren Vergütung gefördert, sondern lediglich mit der Chance auf einen Gewinn.

Konkrete Hinweise auf die Bevorzugung positiver Bewertungen und dementsprechend das Problem der Asymmetrie haben die Ermittlungen im Zusammenhang mit Anreizen nicht ergeben. Theoretisch ist jedoch auch hier – wie bei den Anstupsern – eine entsprechende Ausgestaltung denkbar, die das Gesamtbild der Bewertungen verzerrt.

Händler und Dienstleister bieten ihren Kunden schließlich zum Teil auch offline Anreize wie Gutscheine oder Ermäßigungen in Form eines Freigetränks o. ä., wenn diese dafür auf einem bestimmten Portal eine (positive) Bewertung abgeben. Soweit dabei Einfluss auf den Inhalt der Bewertungen genommen wird, besteht auch hier das Problem der Asymmetrie.

D.1.4. Produkttests

Die oben genannten Anreize werden für Bewertungen gewährt, die der Nutzer für ein von ihm aufgrund seiner eigenen Bedürfnisse ausgewähltes Produkt abgibt. Daneben versuchen Portale und insbesondere auch Händler teilweise gezielt, für von ihnen ausgewählte Produkte Bewertungen zu generieren, indem sie Nutzer als Produkttester dafür gewinnen.

Von den befragten Portalen haben vier angegeben, selbst Produkttester für Rezensionen einzusetzen. Bei Amazon erfolgt dies im Rahmen von Amazon Vine, einem Programm, das in Deutschland zunächst nur ausgewählten Herstellern zur Verfügung stand, seit Dezember 2019 aber Händlern, die die Teilnahmebedingungen erfüllen, offen steht. Dabei erhalten Nutzer, die bereits viele und von anderen Nutzern als hilfreich bewertete Rezensionen für bestimmte Produkte verfasst haben, die Möglichkeit, bestimmte Produkte kostenlos zu testen und dann zu bewerten. Die im Rahmen von Vine verfassten Rezensionen werden von Amazon

³² BGH, Urt. v. 04.07.2002, Az. I ZR 38/00 – *Zugabenbündel*.

entsprechend gekennzeichnet. Bei den anderen Portalen haben Nutzer die Möglichkeit, sich für Produkttests zu bewerben und werden dann vom Portal ausgewählt.

Die Tests erfolgen auch hier in Kooperation mit den Herstellern der getesteten Produkte. Ein Portal hat darüber hinaus angegeben, Gewinnspiele für die Teilnahme an einem Produkttest durchzuführen: Dabei können sich Nutzer des Portals zu einem Gewinnspiel anmelden. Diejenigen, die als Gewinner ermittelt werden, bekommen ein Produkt über einen Zeitraum von 14 Tagen kostenlos zur Verfügung gestellt. Innerhalb dieser Zeit muss der Kunde eine Produktbewertung mit einer gewissen Mindestlänge abgeben und darf dafür dann das Produkt behalten.

Zur Bewertung von Händlern haben die Vertreter von Google berichtet, dass im Rahmen des „Test Shopping Program“ bei ausgewählten Händlern Testkäufe durchgeführt werden, wobei deren Lieferqualität und Service bewertet wird. Die dadurch erzielten Bewertungen fließen mit in die Durchschnittsbewertung für einen Händler ein.

Neben den Portalen selbst initiieren insbesondere auch Händler Produkttests entweder direkt über (soziale) Netzwerke oder Bewertungsvermittler (s. Kapitel E.2.2.). Die dadurch entstehenden Bewertungen sind dann teilweise auch auf den großen Plattformen zu sehen.

D.1.5. Wesentliche Erkenntnisse

Hinsichtlich der Frage, welche Items auf einem Portal bewertet werden können, haben die Ermittlungen ergeben, dass teilweise das Portal selbst, die betroffenen Unternehmen oder auch die Nutzer darüber entscheiden, welche Leistungen überhaupt dargestellt werden. Während die Darstellung einer großen Zahl von (bewerteten) Leistungen grundsätzlich für das Portal und auch für die Verbraucher vorteilhaft ist, können die betreffenden Anbieter der Leistungen von der ggfs. unfreiwilligen Aufnahme und möglicherweise unvorteilhaften Darstellung auf einem Portal durchaus negativ betroffen sein. Bei dieser Konstellation handelt es sich jedoch in erster Linie um ein B2B-Thema, das bereits Gegenstand zahlreicher gerichtlicher Auseinandersetzungen war und wozu bereits eine umfassende Rechtsprechung existiert. Indirekt könnten allerdings auch Verbraucher betroffen sein, wenn diese nicht einschätzen können, warum welche Anbieter auf einem Portal zu finden sind bzw. wie die Darstellung eines Anbieters zustande gekommen ist. Um dies zu vermeiden, sollten Portale die Verbraucher grundsätzlich darüber informieren, wie die Items erfasst werden und insbesondere darüber aufklären, wenn hervorgehobene Darstellungen auf einem kostenpflichtigen Service des Portals beruhen.

Die Voraussetzungen, die Nutzer erfüllen müssen, um Bewertungen abgeben zu können, sind bei den Portalen unterschiedlich streng. Für reine Bewertungs-Leser ist dabei in der Regel nicht erkennbar, welche Bedingungen Verfasser erfüllen müssen. Grundsätzlich stehen die Portale bei der Gestaltung möglicher Einschränkungen zur Abgabe von Bewertungen vor einer Abwägung: Wenn alle Nutzer ohne Einschränkungen eine Bewertung abgeben können, erhält ein Portal mehr Bewertungen, die überdies repräsentativer sein können. Missbrauch lässt sich in einem solchen System allerdings schwieriger kontrollieren. Gibt es umgekehrt Einschränkungen bei der Abgabe wie z. B. bei einem geschlossenen System, in dem nur nach Aufforderung eine

Bewertung abgegeben werden kann, erhält ein Portal weniger Bewertungen, die – je nach Einzelfall – auch weniger repräsentativ sein können (s. Kapitel F.7.3.). Missbrauch lässt sich aber einfacher kontrollieren. Für welches System sich ein Portal entscheidet, hängt auch vom Geschäftsmodell und den tatsächlichen Möglichkeiten des jeweiligen Portals ab: Für ein Handelsportal ist es relativ einfach, Bewertungen nur für den Fall einer Transaktion zuzulassen, während z. B. ein Ärztebewertungsportal keine Möglichkeit hat, einen vergleichbaren Nachweis eines tatsächlichen Arztbesuches zur Voraussetzung für die Abgabe einer Bewertung zu machen.

Anstupser und Anreize, die die Nutzer zur Abgabe einer Bewertung motivieren sollen, sind nach den Erkenntnissen aus der Sektoruntersuchung weit verbreitet und grundsätzlich nicht zu beanstanden. Für die Portale ist es allerdings unbefriedigend, dass keine klar vorgegebene Grenze für zulässige Vergütungen existiert, insbesondere, da die Portale – wie die Ermittlungen gezeigt haben – im Wesentlichen sehr bemüht sind, bei der Gewährung von Anreizen innerhalb eines rechtlich korrekten Rahmens zu bleiben (s. Kapitel F.1.).

Anstupser oder Anreize, die eine asymmetrische Komponente beinhalten dergestalt, dass positive Bewertungen mit einer größeren Wahrscheinlichkeit abgegeben werden als negative, stellen dagegen für Verbraucher ein Problem dar, weil sie dadurch möglicherweise einen falschen Eindruck von den Qualitäten des bewerteten Items erhalten (s. Kapitel F.5.). Die Ermittlungen haben keine Hinweise auf asymmetrische Anreize (Vergütung nur bei positiver Bewertung), teilweise aber auf asymmetrische Anstupser (Abgabe positiver Bewertungen ist einfacher) ergeben.

Produkttests, bei denen Nutzer gezielt motiviert werden, für bestimmte Produkte Bewertungen abzugeben, werden sowohl von einigen Portalen als auch von zahlreichen Händlern eingesetzt. Sofern derartige Produkttests zu unbeeinflussten, aussagekräftigen Bewertungen führen, sind sie grundsätzlich hilfreich für Verbraucher. Allerdings können derartige Produkttest-Bewertungen im Rahmen eines Bewertungssystems unzulässig sein, wenn aufgrund mangelnder Kennzeichnung für den Verbraucher nicht verständlich ist, unter welchen Bedingungen entsprechende Bewertungen zustande gekommen sind (s. Kapitel F.2.).

D.2. Filterung der Bewertungen

In aller Regel durchlaufen abgegebene Bewertungen, bevor sie auf dem jeweiligen Portal veröffentlicht werden, einen Filterprozess.

Ziel des Filterprozesses der Portale ist es, diejenigen Bewertungen zu identifizieren, die nicht dargestellt werden sollen. Dabei geht es schwerpunktmäßig um den Text der Bewertung, also die Rezension, deren Inhalt z. B. aus rechtlichen Gründen nicht veröffentlicht werden darf. Die reine Vergabe einer bestimmten Sterne- oder Punktezahl kann und muss bezüglich ihres Inhalts nicht gefiltert werden, in Frage kommt dabei (nur) eine Prüfung auf Authentizität. Entsprechend führen diejenigen drei der befragten Portale, bei denen die Nutzer im Rahmen ihrer Bewertung keinen Text eingeben, sondern nur Sterne bzw. Smileys vergeben können, keinen Filterprozess durch. Es handelt sich dabei um zwei Portale, auf denen Filme bewertet werden können, sowie Verkäuferbewertungen bei Ebay-Kleinanzeigen. Von den übrigen Portalen, auf denen Nutzer bei einer Bewertung neben einer Sterne- oder Punktevergabe auch eine Rezension verfassen können oder

sogar müssen, hat nur eines angegeben, überhaupt keine Prüfung eingehender Bewertungen vorzunehmen, sondern erst auf explizite Anfrage eines Bewerteten die für diesen abgegebene Rezension zu prüfen.

Mit ca. 80 % veröffentlicht die überwiegende Mehrheit der befragten Portale auf ihrer Seite Bewertungsrichtlinien, manchmal auch als Verhaltensregeln oder Netiquette bezeichnet, die maßgeblich für den internen Filterprozess durch das Portal sind. Teilweise erscheinen diese Richtlinien direkt bei Eingabe einer Bewertung oder als Link, sodass sie für den Nutzer einfach zu finden sind. In weiten Teilen überschneiden sich die Inhalte der Bewertungsrichtlinien der verschiedenen Portale: Es wird um sachliche Ausdrucksweise und einen respektvollen Umgang miteinander gebeten. Die Bewertungen sollen sich auf eigene Erfahrungen mit dem bewerteten Item beziehen; weder soll also über die Erfahrungen Dritter berichtet werden, noch sollen sie sich auf Faktoren oder Erlebnisse beziehen, die nur mittelbar mit dem Item zusammenhängen. Bei der Bewertung eines Produktes soll z. B. keine Rolle spielen, ob der Versand reibungslos funktioniert hat. Schimpfworte, Hassreden, Diskriminierungen und explizite sexuelle Inhalte werden ausgeschlossen. Das Gleiche gilt für Werbung sowie für illegale Inhalte wie Erpressungen u. Ä. Bei vielen Portalen wie z. B. bei Amazon und Jameda umfasst das Werbeverbot explizit auch das Verbot von Produkttests. So heißt es in den entsprechenden Richtlinien „Zur Wahrung der Integrität der Community-Inhalte sind Inhalte und Aktivitäten zu Werbezwecken (direkt oder indirekt) verboten, einschließlich [...] Erstellen, Ändern oder Veröffentlichen von Inhalten im Austausch für jegliche Art von Vergütung (einschließlich kostenfreie oder vergünstigte Produkte) oder im Auftrag einer anderen Person.“³³ Es wird ferner darauf hingewiesen, dass Rezensionen von Produktherstellern, die sich als unvoreingenommene Käufer ausgeben, sowie Rezensionen als Gegenleistung für eine finanzielle Verfügung entfernt würden.³⁴ Manche Vorgaben der Portale ergeben sich aus Besonderheiten des jeweiligen Bewertungssystems bzw. der bewerteten Items. So muss offenbar im Reisebereich besonders deutlich darauf hingewiesen werden, dass es unzulässig ist, mit einer schlechten Bewertung zu drohen, um einen Rabatt auf den Übernachtungspreis zu erhalten.

Bei der konkreten Umsetzung der Bewertungsrichtlinien, also der Ausgestaltung des internen Filterprozesses, zeigten sich zwischen den Befragten große Unterschiede. Dies betrifft den Einsatz einer automatisierten Filterung, die Durchführung einer manuellen Filterung, die nachträgliche Überprüfung von Bewertungen, initiiert durch Nutzer oder Anbieter, und die Konsequenzen des Filterprozesses.

D.2.1. Automatisiertes Filtern

Bei fast einem Viertel der befragten Portale wird jede abgegebene Bewertung vor der Veröffentlichung manuell geprüft. Die Mehrheit der Portale setzt aber einen oder mehrere automatisierte Filter ein, um möglicherweise problematische Rezensionen automatisiert zu identifizieren.

³³ <https://www.amazon.de/gp/help/customer/display.html?nodeId=201929730> (abgerufen am 01.10.2020).

³⁴ https://www.amazon.de/gp/help/customer/display.html/ref=hp_bc_nav?ie=UTF8&nodeId=G3UA5WC5S5UUKB5G (abgerufen am 01.10.2020).

Acht der befragten Portale (13 %) setzen dabei einen Wortfilter ein, der verhindert, dass eine Bewertung bei Verwendung bestimmter Wörter oder Ausdrücke überhaupt abgegeben werden kann. Dem entspricht z. B. das Vorgehen des Dienstleisters Bazaarvoice, so dass die Online-Händler, die deren System verwenden, über eine entsprechende Abgabesperre verfügen. Auch Facebook nutzt einen derartigen „Phrase Blocker“, der automatisch verhindert, dass Beiträge, die bestimmte Wörter oder Ausdrücke enthalten, auf Facebook veröffentlicht werden. Weitere 46 % der Befragten setzen einen Wortfilter ein, der *nach* Abgabe einer Bewertung bestimmte Rezensionen identifiziert. Darüber hinaus oder alternativ verwenden 41 % der Befragten einen weiteren Filter, der etwas komplexere Prüfungen durchführt („Glaubwürdigkeitsfilter“). Damit sind alle Filter umfasst, die über die reine Filterung bestimmter Suchwörter hinaus weitergehende Prüfungen auf Besonderheiten der Bewertung durchführen und z. B. die Herkunft einer Rezension (Geräte-ID, IP-Adresse) berücksichtigen. Die von den automatisierten Filtern identifizierten Bewertungen werden teilweise direkt gelöscht, in den meisten Fällen aber lediglich gefiltert und einer anschließenden manuellen Einzelfallprüfung zugeführt.

Wortfilter sind technisch relativ einfach umzusetzen und werden deshalb von vielen Portalen eingesetzt. Bei den Kriterien, nach denen Bewertungen grundsätzlich gefiltert werden, gab es im Detail einige Unterschiede, im Großen und Ganzen aber auch viele Übereinstimmungen. Wortfilter, die nur deutsche Ausdrücke filtern, enthalten in der Regel um die 300 Wörter. Bei international tätigen Portalen wird teilweise ein Wortfilter für Bewertungen in verschiedenen Sprachen verwendet. Der Filter kann dann leicht mehrere tausend Wörter bzw. Ausdrücke umfassen.

Die folgenden Filterkategorien wurden von den Portalen regelmäßig genannt:

- Schimpfworte, Schmähkritik, Diskriminierungen: Alle hier erfassten Wortfilter filtern nach verschiedenen Arten von Schimpfwörtern, drei Viertel der Filter prüfen außerdem auf diskriminierende Ausdrücke. Genannt wurden in diesem Zusammenhang unterschiedliche Formen von unangebrachter Sprache wie rassistische, sexistische und extremistische Ausdrücke sowie die Verwendung von Emojis, die Schimpfwörter darstellen.
- Hinweise auf die Identität des Verfassers oder anderer Personen: Die am zweithäufigsten verwendete Filterkategorie umfasst alle möglichen Informationen, die Rückschlüsse auf die Identität einer Person zulassen. Das kann sowohl den Verfasser selbst betreffen als auch z. B. Angestellte in dem bewerteten Restaurant u. Ä. Die Portale nutzen dabei unterschiedliche Mittel, um Klarnamen, Telefonnummern, Emailadressen, Kreditkartendaten und Bankverbindungen zu identifizieren und dann zu entfernen.
- Gefahrhinweise: Von zehn der befragten Portale und damit als dritthäufigste Kategorie wurde angegeben, dass ihr Filter Wörter identifiziere, die Hinweise auf von den Produkten ausgehenden Gefahren geben könnten, wie z. B. „Allergie“, „Pusteln“ oder „Splitter“.
- Versteckte Werbung: Als weitere Kategorie, angegeben von acht der Befragten, lassen sich alle möglichen Hinweise zusammenfassen, die Wettbewerber benennen oder auf andere Art und Weise für andere Marktteilnehmer werben. Derartige Hinweise versuchen die Portale zu entfernen, indem ihr Filter die Namen von Wettbewerbern enthält und indem Links auf andere Websites und insbesondere Youtube-Videos aufgespürt werden.

- Sachfremde Bewertung: Einige Portale haben schließlich angegeben, dass sie Begriffe wie „Service“ oder „Lieferung“ als Filter für Produktbewertungen einsetzen. Bei der anschließenden manuellen Prüfung bestehe dann die Möglichkeit, eine Bewertung für ein Produkt zu löschen oder zu verschieben, wenn die Rezension z. B. nur die Qualität der Lieferung betrifft und nicht das Produkt als solches.

Analog zum Wortfilter gibt es auch Portale, die Bilder und andere Medien in Bewertungen automatisiert filtern um z. B. Nacktbilder identifizieren zu können oder Videos, bei denen Urheberrechtsverstöße vorliegen können.

Über einen reinen Wortfilter hinaus verwenden einige Portale weitere Filter, die eine komplexere Prüfung vornehmen. In der Befragung des Bundeskartellamts wurden diese Filter als Glaubwürdigkeitsfilter bezeichnet, weil ein Schwerpunkt dieser Filter darin besteht, die Authentizität abgegebener Bewertungen zu prüfen, verdeckte Produkttests zu identifizieren oder in sonstiger Weise nicht regelkonform entstandene Bewertungen aufzuspüren. Eine strikte Aufteilung aller automatisierten Filterprozesse in den Unternehmen in die beiden Kategorien Wort- und Glaubwürdigkeitsfilter würde allerdings der Vielfalt der Formen, mit denen die Befragten Prüfungen vornehmen, nicht gerecht. Aus den Antworten hat sich vielmehr ergeben, dass Unternehmen mit unterschiedlichen technischen Mitteln, unterschiedlichen Gesamtkonzepten und in unterschiedlichem Ausmaß Filtermaßnahmen durchführen. Auch die Kriterien, die in Bezug auf die Glaubwürdigkeit verwendet werden können, unterscheiden sich bei den verschiedenen Typen von Portalen. Einzelne Portale aus dem Reisebereich führen eine automatische Plausibilitätsprüfung der einzelnen, zusammen mit der Bewertung erfassten Daten durch (z. B. Datum und Ort der Reise und der Bewertungsabgabe). Manche Portale verwenden auch auffällige Verhaltensmuster von Nutzer-Accounts, um Verdachtsfälle zu identifizieren. Was auffällig ist, hängt dabei vom einzelnen Portal ab und davon, wie es typischerweise genutzt wird. Mehrere Portale haben dargestellt, dass sie Machine Learning einsetzen, um Bewertungen zu filtern, die gegen ihre Richtlinien für Nutzerbewertungen verstoßen. Bazaarvoice hatte verschiedene Möglichkeiten dargestellt, nicht-authentische Bewertungen oder verdeckte Produkttests zu entdecken. Dazu gehört u. a. die Analyse von Metadaten wie z. B. die Geschwindigkeit der Bewertungsabgabe, plötzliche Häufungen von Bewertungen, die Region des Verfassers oder Auffälligkeiten bei den IP-Adressen und Geräten. Auch Hinweise in Facebook-Gruppen könnten ggfs. verwendet werden.

Bezogen auf Mehrfachbewertungen berichtete ein Bewertungsportal, dass es eine systemische Erkennung für identische Bewertungsinhalte beim gleichen oder bei verschiedenen Dienstleistern verwendet und dann doppelte Inhalte entfernt. Ein anderes Portal gab an, dass bei drei hohen Bewertungen für einen Anbieter von derselben IP-Adresse oder Email am gleichen Tag alle Bewertungen von dieser Adresse ungültig gemacht würden. Jameda verwendet eine SMS-Identifikation, um dem Problem von Mehrfachbewertungen zu begegnen. Bei Abgabe einer Bewertung muss immer eine E-Mail-Adresse angegeben werden. Besteht der Verdacht, dass ein Nutzer mehrere Bewertungen für den gleichen Anbieter von verschiedenen E-Mail-Adressen abgegeben hat, verlangt Jameda darüber hinaus eine SMS-Identifikation. Dabei wird die E-Mail-Adresse über die 2-Wege-Authentifizierung einer Mobilnummer zugeordnet. Letztere stellt eine deutlich höhere Hürde auf, denn eine zweite Mobilnummer ist deutlich aufwändiger zuzulegen als eine E-Mailadresse. Versucht ein Nutzer unter einer neuen E-Mail-Adresse, aber bei der SMS-Verifikation mit der gleichen Mobilfunknummer zu bewerten, wird er gesperrt.

Als Besonderheit verwendet Yelp einen Filter, der Bewertungen ausdrücklich in empfohlene und nicht empfohlene einteilt. Alle Bewertungen sind für den Nutzer zugänglich, aber nur die empfohlenen gehen in die durchschnittliche Bewertung eines Anbieters ein und nur diese werden direkt angezeigt, ohne dass ein weiterer Klick erforderlich ist. Der Anteil der nicht empfohlenen Bewertungen liegt bei etwa einem Viertel. Dieses spezielle Vorgehen war bereits Gegenstand gerichtlicher Auseinandersetzungen; zuletzt wurde es vom Bundesgerichtshof bestätigt.³⁵

Neben den Filtern, die im Wesentlichen das Vorkommen bestimmter Wörter und Ausdrücke oder auffälliges Bewertungsverhalten aufdecken, verwenden einige wenige Portale auch Filter für schlechte Bewertungen. So haben mehrere Portale explizit angegeben, dass alle Bewertungen mit drei bzw. zwei oder weniger Sternen automatisch einer manuellen Prüfung zugeführt würden, im Gegensatz zu Bewertungen mit höherer Sternenzahl, die nur manuell geprüft würden, wenn sie aus anderen Gründen herausgefiltert werden. Eine Plattform, auf der Dienstleister bewertet werden, fordert bei negativen Bewertungen den Bewertungsfasser auf, die zugrundeliegende Geschäftsbeziehung nachzuweisen, nicht so jedoch bei positiven Bewertungen. Ein Filmbewertungsportal hat angegeben, dass sämtliche Bewertungen für einen bestimmten Film nicht mehr angezeigt worden seien, nachdem der Verleiher sich über die große Zahl negativer Bewertung beschwert habe. In allen diesen Fällen führt dieses asymmetrische Vorgehen der Portale dazu, dass die Wahrscheinlichkeit gelöscht zu werden, für eine negative Bewertung höher ist als für eine positive Bewertung.

D.2.2. Manuelles Filtern

Alle Portale, die überhaupt einen Filterprozess für Bewertungen durchführen, führen zumindest einen Teil dieses Prozesses manuell durch. Dabei werden entweder alle Bewertungen, nur Stichproben oder nur die dafür von einem automatischen Filter ausgewählten Bewertungen von Mitarbeitern des Portals bzw. dem damit beauftragten Dienstleister gelesen und dann weiterverarbeitet. Ein Viertel der Befragten hat angegeben, jede Bewertung einzeln manuell zu prüfen, d. h. jede neue Rezension wird von einem Mitarbeiter gelesen oder zumindest überflogen, bevor sie veröffentlicht wird. Darunter waren auch Portale, die viele Bewertungen erhalten, so dass dieses Vorgehen mit einem beachtlichen Aufwand für das betreffende Unternehmen verbunden ist. Begründet wurde das Vorgehen damit, dass die manuelle Prüfung nicht nur dazu diene, problematische Inhalte zu identifizieren, sondern vor allem auch eine Möglichkeit darstelle, von Nutzern Rückmeldungen über Eigenschaft und Qualität der Angebote auf dem Portal zu erhalten. Diese Informationen würden dann zur Verbesserung des Angebotes genutzt.

Die Portale haben in der Befragung nicht nur Angaben dazu gemacht, wie diese manuelle Prüfung im Unternehmen durchgeführt wird, sondern darüber hinaus auch Unterlagen beigebracht, die unternehmensintern zur Prüfung von Bewertungen verwendet werden. Knapp ein Drittel aller befragten Portale hat entsprechende interne Handlungsanweisungen für Mitarbeiter zur manuellen Vorab-Kontrolle von Bewertungen vorgelegt. Die Vorgaben zur internen Prüfung entsprechen dabei grundsätzlich den externen Vorgaben zur

³⁵ BGH, Urt. v. 14.01.2020, Az. VI ZR 496/18 (u.a.).

Bewertungsverfassung, sind aber teilweise noch deutlich detaillierter und präziser bzw. zeigen mit Beispielfällen, Textbausteinen oder Ablaufschemata auf, wie die allgemein gehaltenen Bewertungsrichtlinien im Einzelfall auszulegen und umzusetzen sind. Hinsichtlich der inhaltlichen Breite und Tiefe der internen Vorgaben zur manuellen Prüfung bestehen zwischen den einzelnen Portalen allerdings große Unterschiede.

Eine manuelle Überprüfung der Sterne-Bewertung in Verbindung mit dem Rezensionstext wird nur von wenigen einzelnen Portalen vorgenommen. Diese Überprüfungen beziehen sich in der Regel darauf, ob die Zahl der vergebenen Sterne grundsätzlich zum Rezensionstext passt oder z. B. versehentlich eine Verwechslung mit dem typischen Notensystem erfolgte (1 = beste Bewertung).

Die manuelle Kontrolle der Rezensionstexte umfasst bei vielen Portalen zunächst die – teilweise bereits vorab automatisiert erfolgte – Überprüfung auf bestimmte Wortkategorien wie Schimpfwörter, persönliche Daten, Links oder irrelevante Inhalte. In den internen Unterlagen einiger Portale finden sich darüber hinaus auch Hinweise zum Umgang mit Rechtschreibfehlern, schlechten Formulierungen, fremdsprachlichen Texten oder sehr langen bzw. sehr kurzen Texten. Während einige Portale ihre Mitarbeiter explizit darauf hinweisen, dass die Rezensionen grundsätzlich nicht verändert werden dürfen, behalten andere sich das Recht vor, Inhalte zu bearbeiten, zu ändern oder zu kürzen.

Je nach Inhalt der Rezension enthalten die internen Vorgaben teilweise auch Hinweise zur Weiterleitung der Bewertung an den Kundenservice (bei echten Reklamationen oder fehlendem Produktbezug), an die Rechtsabteilung (bei möglichen Rechtsverstößen) oder zur Kontaktaufnahme mit dem Bewertungsverfasser (zwecks Problemlösung und ggfs. Überarbeitung der Rezension). In einigen Fällen werden die herausgefilterten Bewertungen zum Zwecke der internen Auswertung mit verschiedenen Codes versehen.

Einzelne Portale, z. B. Holidaycheck oder Tripadvisor, untersuchen im Rahmen der manuellen Überprüfung auch die Authentizität und Glaubwürdigkeit einer Bewertung. Dabei werden – teilweise auf der Basis automatisiert erfasster Auffälligkeiten – beispielsweise das tatsächliche Bestehen einer Kundenbeziehung oder aber das Vorliegen eines Interessenskonflikts (Wettbewerber, Mitarbeiter) kontrolliert, mehrfache Bewertungen durch denselben Verfasser oder für dieselbe Leistung innerhalb eines kurzen Zeitraumes gelöscht sowie Unstimmigkeiten in Bezug auf die IP-Adresse, das Verhaltensmuster oder die Bewertungshistorie eines Nutzers überprüft.

D.2.3. Nutzerinitiierte nachträgliche Überprüfungen

Eine ganze Reihe von Portalen bietet Nutzern die Möglichkeit, Rückmeldungen zu bereits erschienenen Bewertungen geben. Dabei gibt es im Wesentlichen zwei Arten von Rückmeldungen: Zum einen kann eine Bewertung als besonders hilfreich bewertet werden. Häufig geschieht dies durch das Anklicken eines Daumenhoch-Symbols. Bewertungen, die von anderen Nutzern öfter als hilfreich bezeichnet wurden, werden dann teilweise bei der Default-Anzeige der Bewertungen weiter oben angezeigt, teilweise haben Nutzer auch die Möglichkeit, sich gezielt als besonders hilfreich gekennzeichnete Bewertungen anzeigen zu lassen. Zum anderen kann eine Bewertung auch als kritisch gemeldet werden. Häufig geschieht dies durch Anklicken eines

Fähnchen-Symbols aufgrund der englischen Bezeichnung „to flag a review“ für „einen Beitrag melden“. Die folgende Abbildung zeigt für beides ein Beispiel.

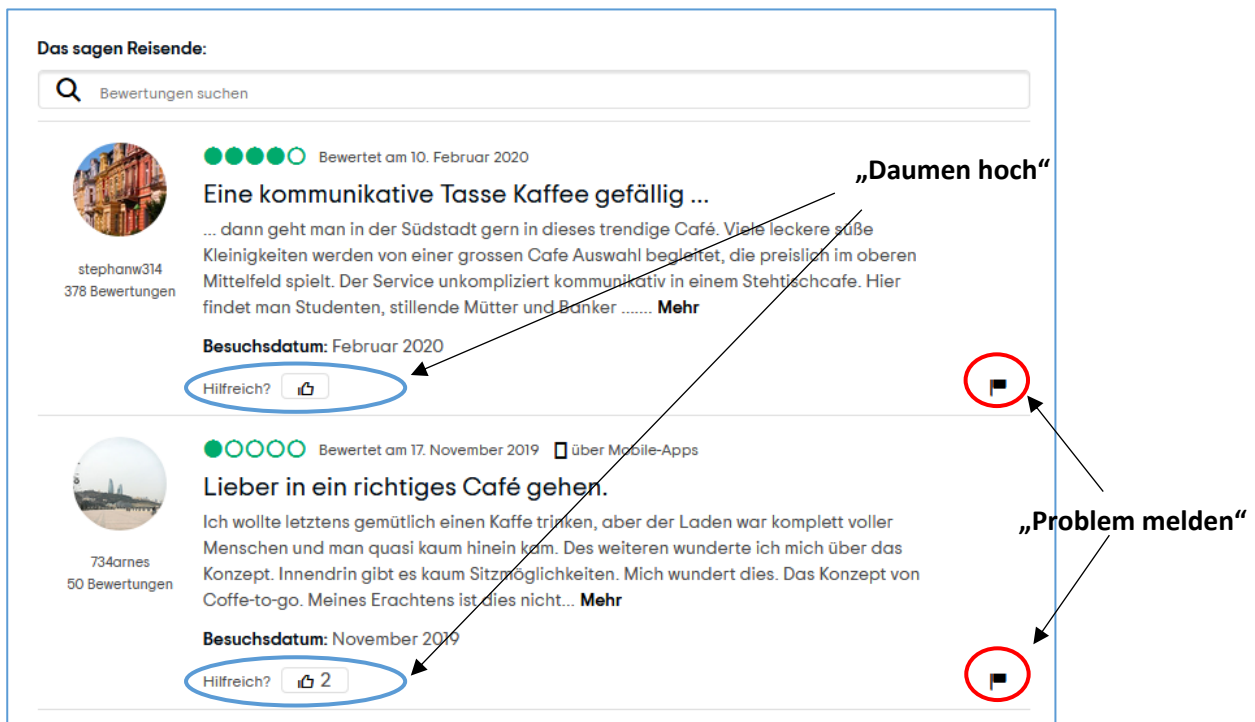


Abbildung 1: Melden von Bewertungen bei Tripadvisor³⁶

Wie hilfreich Bewertungen sind, wird im Folgenden nicht weiter untersucht. Auch eine nicht hilfreiche Bewertung stellt keinen Verbraucherrechtsverstoß dar. Das kann anders sein bei Bewertungen, die z. B. nicht authentisch sind bzw. gegen die Bewertungsrichtlinien des jeweiligen Portals verstoßen. Die Möglichkeit für Nutzer des Portals, Bewertungen als kritisch zu melden, kann solche Fälle aufzudecken helfen. Von den Befragten bieten fast die Hälfte aller Portale im Bereich der Produktbewertungen und über 80 % im Bereich der Unternehmensbewertungen eine solche Möglichkeit an. Die entsprechende Funktion hat Bezeichnungen wie „Beitrag melden“, „Bewertung melden“, „Problem melden“, „Petzen“ oder „Unzumutbar?“.

Bei Google lautet die Bezeichnung bei den allgemeinen Suchergebnissen und auf Google Maps „als unangemessen melden“, das entsprechende Fähnchen erscheint als Symbol allerdings erst, wenn der Nutzer mit der Maus über die entsprechende Stelle fährt. Beim Anklicken öffnet sich ein Fenster, in dem der Nutzer gebeten wird, seine Emailadresse anzugeben sowie eine von vier möglichen Optionen zur Art des Verstoßes anzuklicken. Diese sind:

- (1) Dieser Beitrag enthält Inhalte, die zu Hass aufrufen, gewalttätig oder unangemessen sind.
- (2) Dieser Beitrag enthält Werbung oder Spam.
- (3) Dieser Beitrag weicht vom Thema ab.
- (4) Dieser Beitrag enthält Interessenkonflikte.

³⁶ https://www.tripadvisor.de/Restaurant_Review-g187370-d7599393-Reviews-Black_Coffee_Pharmacy-Bonn_North_Rhine_Westphalia.html (abgerufen am 01.10.2020).

Ein derartiges System nutzen auch einige wenige andere Portale.

Der Anteil der Bewertungen, die von den Nutzern als kritisch gemeldet wird, liegt bei den Portalen, die dazu Zahlen liefern konnten, bei nahezu 0 bis ca. 10 %. Etwas höhere Anteile gab es im Bereich der Produktbewertungen. Im Durchschnitt werden nach den Angaben der Portale 4,8 % aller veröffentlichten Produkt- und 2,9 % aller veröffentlichten Unternehmensbewertungen als kritisch gemeldet. Auch bei den App-Stores gibt es die Möglichkeit, Bewertungen als kritisch zu melden, sie wird aber wesentlich seltener genutzt: Bei beiden großen App-Store-Betreibern ist der Anteil der als kritisch gemeldeten Bewertungen vernachlässigbar. Dies zeigt, dass Nutzer Bewertungen in App-Stores offenbar anders wahrnehmen und anders mit ihnen umgehen als auf Websites im Internet. Gemäß der internen Unterlagen führen einige Portale im Falle der Meldung einer Bewertung eine manuelle Überprüfung verschiedener Inhalts- und Glaubwürdigkeitskriterien durch (s. o.) und nehmen hierzu ggfs. auch Kontakt mit dem Eingebener der Meldung, dem Bewertungsverfasser und/oder dem Bewerteten auf. Während der Überprüfung einer gemeldeten Bewertung wird diese teilweise deaktiviert oder besonders gekennzeichnet.

Von den derart nachträglich gemeldeten Bewertungen werden im Mittel etwa zwei Drittel nach Prüfung unverändert dargestellt, d. h. die gemeldete Bewertung hat sich bei der Prüfung durch das Portal als unkritisch erwiesen. Auffallend ist hier die große Spannweite der Anteile, die unverändert dargestellt werden. Sie reicht von 0 % bei einem Portal, das alle gemeldeten Bewertungen löscht, bis zu 98 % bei einem Portal, das dementsprechend nur 2 % aller gemeldeten Bewertungen nach Prüfung tatsächlich löscht.

D.2.4. Anbieterinitiierte nachträgliche Überprüfungen

Neben Nutzern eines Portals haben auch die dort bewerteten Anbieter die Möglichkeit, an der gleichen Stelle wie die Nutzer eine Bewertung als kritisch zu melden. Häufig erfolgt eine Meldung von Seiten eines Bewerteten allerdings auf einem separaten Weg, z. B. über eine E-Mail an den Kundenservice des Portals. Nur wenige der Befragten erfassen die eingehenden Meldungen statistisch, was ein Indiz dafür sein könnte, dass sie bei vielen Portalen keine große Bedeutung haben. Bei den Portalen, die Meldungen von Anbietern statistisch erfassen, betreffen diese im Bereich Produktbewertungen nur etwa 0,1 % aller Bewertungen. Im Bereich Unternehmensbewertungen sind es mit etwa 1 % aller Bewertungen deutlich mehr. Innerhalb der Gruppe der befragten Portale mit Unternehmensbewertungen zeigten sich dabei auch große Unterschiede, was neben der Auswertung der Antworten in der Befragung auch eine Analyse der mit dem Thema befassten Gerichtsurteile zeigt: Viele der befragten Portale hatten in den letzten Jahren keine Rechtsstreitigkeiten bezogen auf die Zulässigkeit von Bewertungen, andere führen jährlich eine zweistellige Zahl von Verfahren. Übereinstimmend ergibt sich, dass mit Abstand die häufigsten Beschwerden unliebsame Bewertungen der eigenen Leistung betreffen und von Ärzten bzw. Zahnärzten stammen, gefolgt von Anwälten und anderen Dienstleistern wie Autohändlern, Fitnessstudios, Friseuren, Hotels, Immobilienmaklern, Restaurants u.v.a.m. Entsprechend hoch ist mit durchschnittlich über 5 % der Anteil der nachträglich von den bewerteten Anbietern als kritisch gemeldeten Bewertungen bei den beiden befragten Ärztebewertungsportalen Docinsider und Jameda. Google Search und Amazon sind im Bereich Unternehmensbewertungen ebenfalls vergleichsweise häufig von nachträglichen Beschwerden gegen Bewertungen betroffen.

Die unterschiedliche Verteilung der nachträglichen Beschwerden ist sicherlich zum einen darauf zurückzuführen, dass die typischen Bewertungsportale für Ärzte, Anwälte, Restaurants etc. zumeist keine direkte Vermittlung vornehmen und dementsprechend in der Regel offene Bewertungssysteme betreiben (müssen), bei denen der Bewertungsverfasser die tatsächliche Inanspruchnahme der betreffenden Leistung ex ante nicht nachzuweisen braucht. Dementsprechend können die betroffenen Anbieter schneller den Verdacht bekommen, dass die unliebsame Bewertung gar nicht von einem echten Kunden stammt, also nicht authentisch ist. Hinzu kommt, dass es in diesen Bereichen um sehr heterogene, persönliche und emotionale Leistungen geht, deren tatsächliche Qualität oft nicht so sachlich beurteilt werden kann wie beispielsweise bei einem nicht funktionierenden Fön. Und schließlich handelt es sich bei den betroffenen Anbietern zumeist um Berufsgruppen, die an Bewertungen ihrer Person und ihrer Leistungen insgesamt und vor allem im Internet (noch) weniger gewöhnt sind als etwa ein Online-Händler.

Anlass für eine anbieterinitiierte nachträgliche Überprüfung einer Bewertung ist in der Regel, dass ein Anbieter sich gegen eine negative Bewertung zur Wehr setzt. Üblicherweise trägt der Anbieter dabei vor, dass die für ihn unliebsame Bewertung entweder gar nicht auf einer tatsächlichen Inanspruchnahme seiner Leistung beruhe, also nicht authentisch sei, oder dass die Bewertung inhaltlich nicht zutreffend sei bzw. ihn aufgrund der Wortwahl in seinen Persönlichkeitsrechten beschränke.

Zunächst mag es naheliegend erscheinen, dass sich der Anbieter in einem solchen Fall direkt an den Verfasser der Bewertung wendet. Das ist jedoch regelmäßig nicht möglich, da die Identität des Verfassers einer Bewertung typischerweise nicht bekannt ist, denn diese ist nach § 13 Abs. 6 TMG³⁷ geschützt. Die Vorschrift gibt sog. „Diensteanbietern“ vor, die Nutzung von Telemedien in der Außenkommunikation der verschiedenen Nutzergruppen einer Plattform anonym oder unter Pseudonym zu ermöglichen. In der hier betrachteten Konstellation treten zwar die Anbieter nicht anonym auf, in aller Regel aber die Verfasser von Bewertungen. Der Anbieter hat demnach üblicherweise keine Möglichkeit, den Verfasser direkt zu kontaktieren und muss sich folglich an das Portal, auf dem die Bewertung veröffentlicht ist, wenden, um beispielsweise einen Anspruch auf Unterlassung bezogen auf eine unliebsame Bewertung durchzusetzen.³⁸ In einer Stellungnahme eines Rechtsanwalts zum Konsultationspapier wurde angeregt, die Rechtsdurchsetzung diesbezüglich durch eine Ergänzung des § 14 Abs. 3 TMG zu vereinfachen. Nach § 14 Abs. 3 TMG darf ein Diensteanbieter im Einzelfall Auskunft über bei ihm vorhandene Bestandsdaten erteilen, soweit dies zur Durchsetzung zivilrechtlicher Ansprüche wegen der Verletzung absolut geschützter Rechte aufgrund rechtswidriger Inhalte, die von § 1 Absatz 3 des Netzwerkdurchsetzungsgesetzes³⁹ erfasst werden, erforderlich ist. Hier sollte – so die Anregung in der Stellungnahme – ergänzt werden, dass dies auch für Diensteanbieter gelte, die nicht in den Anwendungsbereich des Netzwerkdurchsetzungsgesetzes fallen.

³⁷ Telemediengesetz v. 26.02.2007 (BGBl. I S. 179), zuletzt geändert durch Art. 11 des Gesetzes v. 11.07.2019 (BGBl. I S. 1066) – TMG.

³⁸ S. dazu auch BGH, Urteil vom 1. März 2016, Az. VI ZR 34/15, Rn. 40 – *jameda.de II*.

³⁹ Netzwerkdurchsetzungsgesetz v. 1. September 2017 (BGBl. I S. 3352), zuletzt geändert durch Artikel 274 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1328) – NetzDG.

Nach geltender Rechtsprechung ist ein Portal für die auf seiner Seite veröffentlichten, rechtswidrigen Inhalte so lange nicht haftbar, wie es von der Rechtswidrigkeit keine Kenntnis hat, sofern es dabei keine Verkehrs- bzw. Verhaltenspflichten verletzt. Das Portal ist also nicht grundsätzlich dazu verpflichtet, im Vorfeld alle Bewertungen auf ihre Rechtmäßigkeit zu prüfen. Eine Verpflichtung ergibt sich aber in jedem Fall, sobald das Portal auf einen möglichen Verstoß hingewiesen wird, also z. B. mit der Meldung des Anbieters, dass bei einer konkreten Bewertung der Verdacht auf einen Verstoß bestehe.⁴⁰ Im Wesentlichen durch ein Urteil des Bundesgerichtshofs bestätigt⁴¹ hat sich ein als „Notice-and-Take-Down“ bezeichneter Prozess im Umgang mit derartigen Meldungen etabliert, bei dem für die Überprüfung der gemeldeten Bewertung ein als Pingpong bezeichnetes Verfahren angewendet wird, das vom Bundesgerichtshof ursprünglich für die Prüfung eines anderen Inhalts im Netz, nämlich eines Blog-Beitrags, vorgeschlagen wurde.⁴²

Zur Frage der Haftung durch das Portal hat der Bundesgerichtshof in seinem Urteil „jameda.de II“ folgendes festgestellt: Ist eine Bewertung durch ein Portal nicht verändert worden, sondern hat sie nur den üblichen Filterprozess durchlaufen, so ist davon auszugehen, dass das Portal sich eine Bewertung nicht zu Eigen gemacht hat. In diesem Fall kommt für das Portal lediglich eine Haftung als „mittelbarer Störer“ in Betracht.⁴³ Grundsätzlich gelte zwar für Diensteanbieter wie hier der Portale das Haftungsprivileg (§ 10 TMG), die Möglichkeit der Geltendmachung eines Unterlassungsanspruchs, der seine Grundlage in einer vorangegangenen Rechtsverletzung findet, werde dadurch aber nicht eingeschränkt.⁴⁴ Die Haftung als mittelbarer Störer darf allerdings – so der Bundesgerichtshof – nicht über Gebühr auf Dritte erstreckt werden, welche die rechtswidrige Beeinträchtigung nicht selbst vorgenommen haben. Die Haftung eines Portals setze deshalb die Verletzung von Verhaltenspflichten, insbesondere von Prüfpflichten, voraus. Deren Umfang bestimme sich danach, ob und inwieweit dem Portal nach den Umständen des Einzelfalls eine Verhinderung der Verletzung zuzumuten sei. Ein Hostprovider sei zur Vermeidung einer Haftung als mittelbarer Störer nicht verpflichtet, die von den Nutzern ins Netz gestellten Beiträge vor der Veröffentlichung auf eventuelle Rechtsverletzungen zu prüfen. Verantwortlichkeit entstehe erst, sobald er Kenntnis von einer möglichen Rechtsverletzung in dem konkreten Einzelfall erlangt. Weist ein Betroffener den Hostprovider auf eine Verletzung seines Persönlichkeitsrechts durch den Nutzer seines Angebots hin, könne der Hostprovider darüber hinaus verpflichtet sein, künftig derartige Störungen zu verhindern.⁴⁵

⁴⁰ S. dazu § 10 TMG, der fast wortgleich aus Art. 14 der Richtlinie 2000/31/EG des Europäischen Parlaments und des Rates v. 08.06.2000 über bestimmte rechtliche Aspekte der Dienste der Informationsgesellschaft, insbesondere des elektronischen Geschäftsverkehrs, im Binnenmarkt ("Richtlinie über den elektronischen Geschäftsverkehr"), Abl. EG Nr. L 178 v. 17.07.2000, S. 1 (kurz E-Commerce-Richtlinie) übernommen wurde, sowie die Erwägungsgründe 42 bis 46 der E-Commerce-Richtlinie.

⁴¹ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, – *jameda.de II*.

⁴² BGH, Urt. v. 25.10.2011, Az. VI ZR 93/10, Rn. 27 – *Blog-Eintrag*.

⁴³ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, Rn. 16 ff. – *jameda.de II*.

⁴⁴ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, Rn. 19 – *jameda.de II*.

⁴⁵ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, Rn. 22, 23 – *jameda.de II*.

Zu welchen konkreten Überprüfungsmaßnahmen der Hostprovider verpflichtet ist, bestimmt sich nach den Umständen des Einzelfalls.⁴⁶ Im vorliegenden Fall ging es darum, substantiiert darzulegen, dass tatsächlich ein Behandlungskontakt zwischen Arzt und Bewertungsverfasser stattgefunden hatte. Das Gericht stellt einerseits dar, dass der Prüfungsaufwand für das Portal angemessen sein muss. So darf im konkreten Fall das Geschäftsmodell des Ärztebewertungsportals Jameda durch die Überprüfung nicht gefährdet werden. Das Gericht hält deshalb hier die rein reaktive Prüfung von Bewertungen des Portals für angemessen. Diese müsse andererseits jedoch sorgfältig und auf eine wirkliche Klärung des Sachverhalts ausgerichtet sein. Aufgrund der Anonymität des Bewertungsverfassers gegenüber dem Anbieter sei der grundsätzlich darlegungs- und beweisbelastete Kläger in seinen Möglichkeiten zur Sachaufklärung stark eingeschränkt. Das Portal treffe daher eine sekundäre Darlegungslast.⁴⁷ Die Frage, inwieweit Portale durch bestehende Verkehrs- bzw. Verhaltenspflichten auch verpflichtet sind, eine Vorab-Prüfung durchzuführen, ist vom Bundesgerichtshof im Urteil „jameda.de II“ nicht behandelt worden. In der Entscheidung ging es ausschließlich um die Verhältnismäßigkeit einer ausführlichen nachträglichen Überprüfung. Das Gerichtsurteil hat weitreichende praktische Folgen: Wenn durch eine entsprechende Mitteilung des Anbieters an das Portal („notice“) im Einzelfall Zweifel an der Authentizität eines Kontaktes geweckt werden, setzen viele Portale sofort und weitgehend standardisiert einen Prozess zur Überprüfung der Bewertung in Gang.

Nach den Ermittlungen wird der recht aufwändige Notice-and-Take-Down-Prozess dabei allerdings auch nicht von allen Portalen eingesetzt. Amazon beschreibt z. B. einen etwas anderen Prozess, der Bewertungen nicht vollständig löscht, sondern u. a. darauf ausgerichtet ist, solche Bewertungen zu identifizieren, bei denen ein Anbieter für Probleme beim Lieferprozess kritisiert wird, die er selbst nicht zu verantworten hat. Diese Bewertungen werden dann mit einem entsprechenden Hinweis von Amazon ergänzt. Auch andere Portale verwenden teils branchenabhängig unterschiedliche Prozesse. Gerade gegenüber Ärzten kommt jedoch in aller Regel der Notice-and-Take-Down-Prozess zum Einsatz. Jameda hat angegeben, das Verfahren in jedem einzelnen Verdachtsfall durchzuführen; dabei handelt es sich um eine fünfstellige Zahl pro Jahr. In der Regel wird dabei zunächst die Bewertung vorübergehend bis zur Klärung des Streitfalls entfernt. Während des Klärungsprozesses, der normalerweise innerhalb weniger Wochen abgeschlossen ist, wird zunächst der Verfasser aufgefordert, einen Nachweis für seinen Kontakt zu liefern bzw. seine Bewertung zu substantiieren. Im Falle einer Ärztebewertung kann ein Patient z. B. durch eine Beschreibung der Praxiseinrichtung oder durch das Einreichen von Unterlagen wie Rechnungen, Rezepte, Terminkarten o. Ä. Indizien dafür liefern, dass tatsächlich ein Behandlungskontakt stattgefunden hat. Reagiert der Verfasser innerhalb einer festgesetzten Frist (bei Jameda: 3 Wochen) nicht, wird die Bewertung entfernt („take down“). Andernfalls leitet das Portal eine ggfs. anonymisierte Fassung der Stellungnahme an den Anbieter weiter, der dann dazu wiederum Stellung nehmen kann. Dieses Pingpong kann sich noch mehrmals wiederholen, bis das Portal die Entscheidung trifft, die Bewertung entweder wieder zugänglich zu machen oder sie endgültig zu entfernen („take down“). Ein Portal berichtete, dass die beanstandete Bewertung sehr häufig im Rahmen des Notice-and-take-down-

⁴⁶ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, Rn. 38 – *jameda.de II*.

⁴⁷ BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15, Rn. 40, 42, 46, 47 – *jameda.de II*.

Prozesses gelöscht würde, in 60 % der Fälle allein deshalb, weil der Verfasser der Bewertung innerhalb der vorgegebenen Frist nicht reagiert.

In den Fällen, in denen ein Portal entscheidet, eine Bewertung beizubehalten, kann der Bewertete noch vor einem Zivilgericht auf Unterlassung nach § 823 Abs. 1, § 1004 Abs. 1 Satz 2 BGB⁴⁸ analog klagen. Auch von diesen Klagen sind Portale, je nachdem, welche Berufsgruppe dort bewertet wird, sehr unterschiedlich häufig betroffen. Die große Mehrzahl der Fälle, die vor Gericht verhandelt werden, betrifft Ärzte; allein Jameda wird mehr als 50 Mal im Jahr verklagt.

Durch das höchstrichterlich bestätigte Vorgehen beim Notice-and-Take-Down-Verfahren mit dem Pingpong-Prinzip besteht derzeit eine gewisse Rechtssicherheit, die es Portalen ermöglicht, die Risiken, die mit der Veröffentlichung von Bewertungen verbunden sind, gut einschätzen zu können. Zwar ist es für ein Portal im Einzelfall immer am einfachsten und günstigsten, im Falle eines Antrags auf Unterlassung die betroffene Bewertung zu entfernen, allerdings würde dies auf Dauer dazu führen, dass ein großer Teil aller unliebsamen und damit in der Regel negativen Bewertungen von einem Portal auf Intervention des Bewerteten entfernt würde. Für den Verbraucher, der das Portal als Informationsquelle nutzt, sind negative Bewertungen jedoch mindestens genauso wichtig wie positive, so dass ein Portal für Verbraucher dann nicht mehr attraktiv wäre. So lässt sich erklären, dass Portale wie Google und Jameda trotz der ihnen dadurch entstehenden Kosten regelmäßig oder sogar immer den recht aufwändigen Notice-and-Take-Down-Prozess durchführen und darüber hinaus Gerichtsverfahren initiieren, um möglichst viele negative, unliebsame, aber rechtlich zulässige Bewertungen beibehalten zu können.

Grundsätzlich birgt das Vorgehen, Nutzerbewertungen allein im Notice-and-Take-Down-Verfahren auf ihre Authentizität zu prüfen, in jedem Fall die Gefahr, dass auf einem Portal mehr negative als positive Bewertungen aussortiert werden. Denn in vielen Fällen scheuen Nutzer den Aufwand, im ersten Schritt des Pingpong-Verfahrens zur Glaubwürdigkeit ihrer Bewertung vorzutragen, für den sie in keiner Weise vergütet werden. Bei entsprechend fehlender Rückmeldung bleibt dem Portal in der Regel keine andere Wahl, als die Bewertung zu entfernen. Die Meldung und anschließende Löschung positiver Bewertungen dürfte im Rahmen eines reinen Notice-and-Take-Down-Verfahrens demgegenüber die Ausnahme sein, da hiergegen allenfalls ein Wettbewerber vorgehen würde, für den der Beleg eines Verstoßes deutlich schwieriger wäre.

D.2.5. Konsequenzen und Sanktionen

Auch in Bezug auf den Umgang mit herausgefilterten Bewertungen und den betroffenen Bewertungsverfassern bzw. Anbietern haben die Antworten der Portale sowie die übersandten internen Unterlagen deutliche Unterschiede zwischen den verschiedenen Portalen ergeben.

⁴⁸ Bürgerliches Gesetzbuch in der Fassung der Bek. v. 02.01.2002 (BGBl. I S. 42, 2909; 2003 I S. 738), zuletzt geändert durch Art. 1 des Gesetzes v. 19.03.2020 (BGBl. I S. 541) – BGB.

Im Durchschnitt über alle befragten Portale werden 7 % der dort abgegebenen Bewertungen aufgrund von Filtermaßnahmen am Ende nicht dargestellt. Dieser Durchschnittswert hat sich im Verlauf der Jahre 2014 bis 2018, für die die Werte abgefragt wurden, kaum verändert. Unternehmensscharf ergeben sich bereits für das Jahr 2018 deutliche Unterschiede, wie die folgende Grafik zum jeweiligen Anteil der nicht veröffentlichten Bewertungen für die einzelnen befragten Portale zeigt.

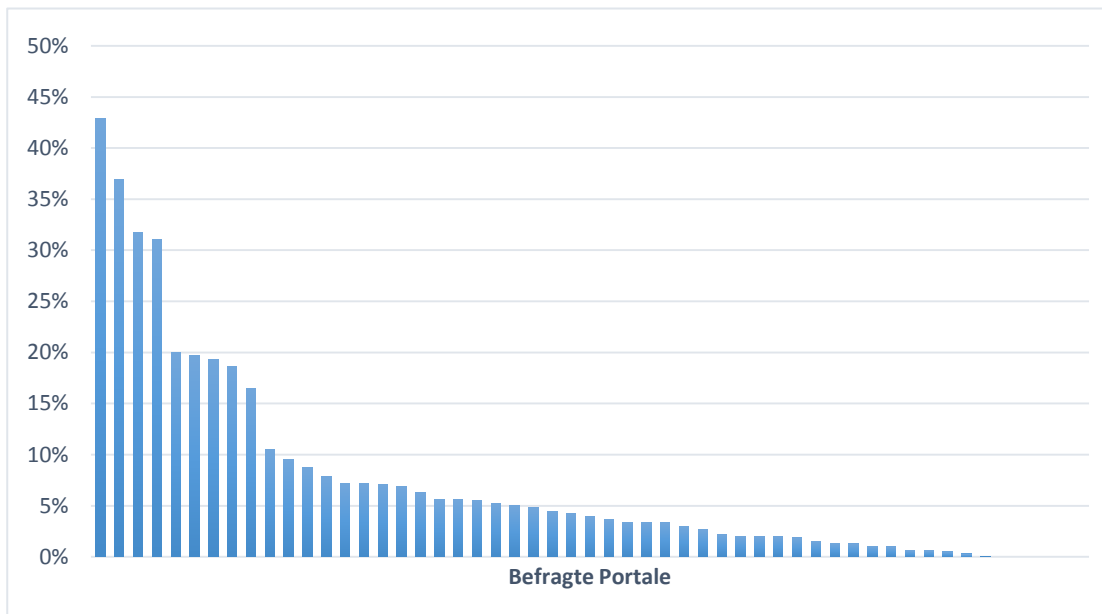


Abbildung 2: Anteil nicht dargestellter Bewertungen bei den Befragten im Jahr 2018

Die Mehrheit der Portale hat angegeben, Rezensionen, die herausgefiltert werden, komplett zu entfernen. Aus den internen Unterlagen der Portale geht hervor, dass einige ausdrücklich keinerlei Veränderungen von Rezensionstexten vornehmen, da sie sich der damit verbundenen Haftungsproblematik sehr bewusst sind. Die übrigen Portale entfernen nur die kritischen Inhalte bzw. ersetzen entsprechende Stellen durch Sternchen. Ein Portal gab z. B. an, bei „wenigen Rezensionen editierende Eingriffe (Kosmetik, Rechtschreibung, Kürzung von Nachnamen)“ vorzunehmen. Andere differenzieren nach dem Grund der Beanstandung: Bewertungen, die Schimpfworte enthalten, werden komplett gelöscht, während bei der Angabe von Klarnamen nur der Namen entfernt, die übrige Rezension aber erhalten bleibt. Google hat z. B. angegeben, dass bei etwa 3 % aller Bewertungen für Google Shopping Informationen entfernt werden, die eine Identifikation des Bewertungsverfassers ermöglichen könnten; die Rezension bleibt dabei aber erhalten.

Wird eine Rezension gefiltert und schließlich nicht angezeigt, könnte die dazugehörige Sternebewertung entweder ebenfalls gelöscht oder dennoch angezeigt werden. Ein Grund für die Entfernung der Sternebewertung mit der Rezension könnte darin bestehen, dass jemand, der z. B. eine Rezension aus sachfremden Motiven verfasst, sich bei der Sternebewertung genauso von sachfremden Motiven leiten lässt. Andererseits kann eine Rezension auch entfernt werden, weil der Nutzer darin z. B. eine Angabe gemacht hat, die auf seine Identität schließen lässt. In diesem Fall wäre es eher angemessen, die Sternebewertung weiterhin anzuzeigen. Die meisten der befragten Portale haben sich dafür entschieden, beim Entfernen einer Rezension auch die damit verbundene Sternebewertung zu entfernen. Nur ein kleiner Teil der Portale, insbesondere im Bereich der Unternehmensbewertungen, hat angegeben, in diesen Fällen die Sternebewertung dennoch, ohne die zugehörige Rezension, anzuzeigen.

Bei manchen Portalen werden die Verfasser der Bewertung teilweise in die Prüfung mit einbezogen, indem sie die Möglichkeit bekommen, Stellung zu nehmen bzw. ihre Bewertung zu ändern oder zumindest darüber informiert werden, dass und ggfs. warum ihre Bewertung nicht veröffentlicht wird. Von den befragten Portalen haben gut ein Drittel im Bereich Produktbewertungen und zwei Drittel im Bereich Unternehmensbewertungen angegeben, dass bei ihnen eine entsprechende Information erfolgt.

Von den aufgrund der Meldung durch einen Nutzer manuell geprüften Bewertungen werden im Bereich der Produktbewertungen bei vielen Plattformen etwa 80 % der gemeldeten Bewertungen nach Prüfung weiterhin unverändert dargestellt. Die große Mehrheit der Plattformen hat hier angegeben, grundsätzlich keine Änderungen an Bewertungen vorzunehmen, so dass die verbleibenden, als kritisch identifizierten Bewertungen gelöscht werden. Ein kleines Portal hat angegeben, dass alle als kritisch gemeldeten Bewertungen gelöscht würden. Im Bereich Unternehmensbewertungen wird nur etwas mehr als die Hälfte aller Bewertungen nach der Prüfung unverändert übernommen. Die meisten Portale nehmen hier – ggfs. in Absprache mit dem Verfasser – teilweise Veränderungen an den Bewertungen vor, das betrifft durchschnittlich knapp 10 % der Bewertungen. Ein knappes Drittel der Bewertungen wird gelöscht bzw. von der Seite genommen. Über die beiden Kategorien Produkt- und Unternehmensbewertungen hinweg zeigt sich, dass bei Portalen, die ein sehr ausgereiftes Bewertungssystem mit gründlichen Vorab-Filtern betreiben, der Anteil an gemeldeten Bewertungen, die nach Prüfung unverändert angezeigt werden können, mit in der Regel über 80 % sehr hoch ist.

In Bezug auf die betroffenen Bewertungsverfasser oder Anbieter setzen einige der befragten Portale Sanktionen ein, um die weitere Abgabe von Bewertungen, die nicht ihren Bewertungsrichtlinien entsprechen, zu verhindern. So gibt es die Möglichkeit, die Schreibrechte eines Nutzers, von dem entsprechende Bewertungen stammen, einzuschränken oder weitergehend dessen Konto vorübergehend oder dauerhaft zu sperren. Plattformen haben darüber hinaus die Möglichkeit, die Konten der Anbieter, für die gehäuft verdächtige Bewertungen abgegeben werden, vorübergehend oder dauerhaft zu sperren.

Die in der Befragung am häufigsten genannte Sanktion bei Auffälligkeiten von Bewertungen oder Verstößen gegen die Bewertungsrichtlinien des Portals war die vorübergehende oder dauerhafte Sperrung des Nutzerkontos. Ein Viertel der Befragten hat angegeben, dass sie tatsächlich im Jahr 2018 entsprechende Kontosperrungen durchgeführt hätten. Etwa die Hälfte dieser Portale hat sowohl vorübergehende als auch dauerhafte Sperrungen vorgenommen, bei der anderen Hälfte kam es nur zu dauerhaften Sperrungen, die auch insgesamt wesentlich häufiger sind. Manche Portale haben darauf hingewiesen, dass bei ihnen Nutzer auch ohne Nutzerkonto bewerten könnten und sie daher nicht über die Möglichkeit verfügen, einen Nutzer zu sperren.

Betroffen von den Sanktionen waren in der Regel weit unter 1 % der Konten. Nur zwei Portale – eines im Bereich Produkt-, eines im Bereich Unternehmensbewertungen – haben angegeben, jeweils gut 3 % aller Nutzerkonten gesperrt zu haben. Dabei handelt es sich um zwei Portale, bei denen nach eigenen Angaben in größerem Umfang automatisiert Nutzerkonten kreiert oder auf fremde Nutzerkonten zugegriffen wird, um über diese gefälschte Bewertungen abzugeben.

Portale, die als Plattformen auf ihrer Seite Angebote Dritter darstellen, haben außerdem grundsätzlich die Möglichkeit, bestimmte Anbieter vorübergehend oder dauerhaft zu sperren. Tatsächlich genutzt wurde diese Sanktionsmöglichkeit allerdings nur von etwa 15 % der befragten Portale sowohl im Bereich der Produkt- als auch der Unternehmensbewertungen. Der zurückhaltende Umgang mit dieser Sanktion mag darin begründet sein, dass die Sanktion für den Anbieter eine erhebliche finanzielle Einbuße bedeuten kann, wenn ein großer Teil seines Umsatzes über die Plattform generiert wird. Eine Hotelbuchungsplattform hat angegeben, dass sie Hotels daher zunächst verwarnt, wenn sie entsprechende Bewertungen für das Hotel findet. Zwei andere Plattformen haben angegeben, dass sie den Eintrag des Anbieters auf ihrer Plattform mit einem Manipulationshinweis versehen. Jameda versieht Profile von Ärzten, bei denen die interne Überprüfung einen Verdacht auf unzulässige Bewertungen ergibt, der sich nicht aufklären lässt, mit einem Warnhinweis: An der Durchschnittsnote ist ein kleines, rotes Ausrufezeichen zu sehen, dem ein längerer Text hinterlegt ist, der darauf hinweist, dass bei einzelnen Bewertungen auf dem Profil Auffälligkeiten festgestellt wurden, die auf Manipulationen hinweisen. Der Warnhinweis habe sich als relativ effektives Mittel erwiesen, Ärzte dazu zu bewegen, Manipulationen aufzuklären und zu beenden. Die Zulässigkeit eines derartigen Hinweises ist teilweise gerichtlich bestätigt,⁴⁹ teilweise abgelehnt⁵⁰ worden. In mehreren Fällen wurde Berufung beim OLG Frankfurt eingelegt, die Entscheidungen des Berufungsgerichts stehen noch aus (Stand: September 2020). Das reine Löschen von Bewertungen ist gerichtlich bestätigt im Fall eines begründeten Verdacht ebenfalls zulässig.⁵¹ Es halte – so Jameda – die Verfasser aber kaum davon ab, weitere Bewertungen zu platzieren.

Eine weitere mögliche Sanktion der Portale besteht darin, gerichtlich gegen diejenigen vorzugehen, die zu löschende Bewertungen verursacht haben. Bei einer Plattform kommen dafür zum einen die Anbieter in Betracht, die für sich selbst (positive) Bewertungen initiiert haben, und zum anderen die Bewertungsvermittler, die den Kontakt mit den Bewertungsverfassern organisieren. Diese Vorgehensweise wird im Kapitel E.3.1. dargestellt und diskutiert.

D.2.6. Wesentliche Erkenntnisse

Über alle Portale hinweg werden etwa 7 % aller Bewertungen aufgrund bestehender Filtermaßnahmen nicht dargestellt. Bezüglich der Sorgfalt und Tiefe der Filtermaßnahmen haben die Ermittlungen erhebliche Unterschiede zwischen den einzelnen Portalen ergeben: Viele Portale prüfen abgegebene Bewertungen zwar automatisiert oder manuell auf beleidigende oder unsachliche Ausdrücke hin, auf Werbung oder auf mögliche Datenschutzverstöße. Eine weitergehende Prüfung auf Authentizität oder das Vorliegen verdeckter Produkttests findet jedoch Großteils gar nicht oder nur cursorisch statt. Dabei zeigen einzelne Portale, dass es durchaus Möglichkeiten gibt, manuell, aber auch unterstützt durch Algorithmen die fehlende Authentizität oder

⁴⁹ LG Frankfurt a. M., Beschl. v. 09.06.2020, 2-03 O 167/20 (nicht veröffentlicht).

⁵⁰ LG Kassel, Beschl. v. 15.06.2020, 10 O 703/20 (nicht veröffentlicht).

⁵¹ OLG München, Urt. v. 27.02.2020, 29 U 2584/19.

das Vorliegen verdeckter Produkttests zumindest teilweise aufzuspüren. Dies geschieht v. a. durch die Analyse von Metadaten der Bewertungsabgabe (z. B. IP-Adresse) oder zusätzlich erfragte Informationen (z. B. Datum des Aufenthalts), über die das Portal verfügt.

Bewertungen in beleidigender Sprache oder mit Werbung für andere Marktteilnehmer schädigen in erster Linie die Betroffenen. Für den Verbraucher schädigend sind dagegen vor allem nicht-authentische oder manipulierte Bewertungen sowie nicht als solche gekennzeichnete Produkttests, die zwar in angemessener Sprache verfasst sind, sich auf das zu bewertende Item beziehen und keine Datenschutzprobleme aufwerfen, denen aber keine tatsächliche Nutzung des bewerteten Items zugrunde liegt oder die nicht frei und unbeeinflusst zustande gekommen sind. Zur Identifikation und ggf. Entfernung nicht-authentischer oder manipulierter Bewertungen und verdeckter Produkttests sind spezifische Filter sowie ggfs. über das Löschen hinausgehende Sanktionen erforderlich. Umso problematischer ist es daher, dass die meisten Portale keine diesbezüglichen Filter verwenden bzw. die technischen Möglichkeiten einer derartigen Filterung nicht ausschöpfen (s. Kapitel F.2., F.3. und F.4.).

Unabhängig von der Art der verwendeten Filter werden auf vielen Portalen Bewertungen nach einer Prüfung entweder unverändert beibehalten oder vollständig gelöscht. Eine Überarbeitung findet dagegen prinzipiell nicht statt. Diese Portale sind sich anscheinend des Haftungsrisikos bewusst, das mit der Veränderung von Bewertungen und damit verbunden einem Zu-Eigen-Machen der Bewertung verbunden sein kann.

Die Möglichkeit für Nutzer und Anbieter, ex post auffällige Bewertungen zu melden, stellt grundsätzlich eine hilfreiche Ergänzung zu einem ex ante durchgeführten Filtersystem dar. Bezogen auf Nutzer gilt dies jedoch nur, wenn diese Möglichkeit klar zu erkennen ist und das Portal für den Umgang mit gemeldeten Bewertungen entsprechende Maßnahmen vorsieht. Je besser das Vorab-Filtersystem, desto seltener sollten auf diesem Wege allerdings noch Missstände aufgedeckt werden. Auch im Streit um für die Anbieter unliebsame Bewertungen sind schon aufgrund der Anonymität der Verfasser die Portale verantwortlich für nachträgliche Überprüfungen. Die überwiegende Mehrheit von Beschwerden über unliebsame Bewertungen stammt von Ärzten bzw. Zahnärzten. Besonders betroffen von anbieterinitiierten nachträglichen Überprüfungen sind daher Portale, auf denen diese bewertet werden. Im gerichtlich bestätigten Notice-and-take-down-Prozess erhalten beide Seiten die Gelegenheit, sich zu dem Sachverhalt zu äußern. Das Portal übernimmt dabei die Vermittlung der (anonymisierten) Erläuterungen an die jeweils andere Seite und entscheidet am Ende über den Verbleib oder die Löschung der Bewertung. Das Verfahren führt in Summe dazu, dass unliebsame (also in der Regel negative) Bewertungen häufig von den Portalen gelöscht werden. (s. Kapitel F.5.).

Kritisch erscheinen schließlich Filtermaßnahmen, automatisiert oder manuell, die dazu führen, dass negative Bewertungen eher als positive Bewertungen von einem Filter erfasst und anschließend nicht dargestellt werden (s. Kapitel F.5.).

D.3. Darstellung der Bewertungen

Ihre eigentliche Wirkung entfalten Bewertungen erst, wenn sie auf einem Portal letztlich dargestellt, vom Verbraucher wahrgenommen und ggfs. bei seiner Konsumentenscheidung berücksichtigt werden. Typischerweise stellen Portale an prominenter Stelle des Angebots eine zusammengefasste (Sterne-)Bewertung dar; durch Anklicken dieser Zusammenfassung oder durch Herunterscrollen auf der Seite lassen sich dann auch einzelne Bewertungen ansehen. Beispielhaft zeigt die folgende Abbildung das Vorgehen eines Online-Händlers: Auf der Produktdetailseite ist direkt unter dem Artikelnamen zu sehen, wie viele Sterne das Produkt in Bewertungen durchschnittlich erhalten hat, dahinter steht die Anzahl der Bewertungen, die dem Durchschnitt zugrunde liegt. Fährt man mit der Maus über die Angabe der Durchschnittsbewertung, öffnet sich ein Fenster, das die Verteilung der Bewertungen zeigt.

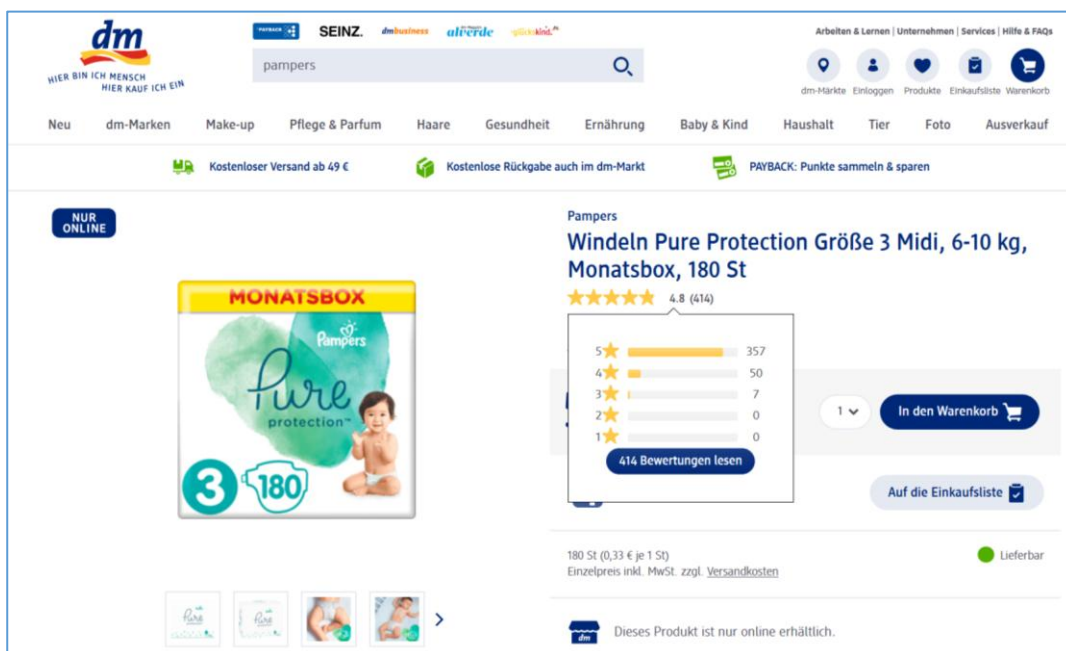


Abbildung 3: Anzeige Bewertungszusammenfassung bei dm⁵²

Klickt man auf die Sterne, öffnet sich eine Seite, auf der die Bewertungen einzeln aufgeführt sind. In dem Beispiel der Abbildung sind die verfügbaren Bewertungen nach „Relevanz“ sortiert:

⁵² <https://www.dm.de/pampers-windeln-pure-protection-groesse-3-midi-6-10-kg-monatsbox-p8001841115917.html> (abgerufen am 01.10.2020).

Bewertungen Jetzt Produkt bewerten

Produktbewertungen, die auf der dm.de abgegeben wurden und der Netiquette & den Nutzungsbedingungen entsprechen, haben an der Verlosung eines Gewinnspiels teilgenommen.
Bewerten Sie Ihre Produkte & nehmen Sie auch am Gewinnspiel teil.

5 ★	382
4 ★	53
3 ★	8
2 ★	0
1 ★	0

1-8 von 443 Bewertungen Sortieren nach: Relevanteste

★★★★★ Katrin · vor einem Jahr
Pampers PureProtection

Wir hatten das große Glück die neue Pampers PureProtection kostenfrei testen zu dürfen. Normalerweise verwenden wir immer die PremiumProtection Windeln und sind damit sehr zufrieden. Wir haben das Gefühl, dass die PureProtection sogar noch hautverträglicher ist, obwohl unser kleiner Schatz nur äußerst selten Rötungen Im Windelbereich aufweist. Die PureProtection Windel hält sehr lange trocken, ist super dünn und hat ein unglaublich niedliches Design. Leider finden wir sie insgesamt etwas teuer, auch wenn Nachhaltigkeit natürlich Ihren Preis haben muss. Insgesamt sind wir super zufrieden und würden die Windel auf jeden Fall weiterempfehlen. Katrin

Ursprünglich erschienen auf pampers.de

★★★★★ Bianca · vor einem Jahr
Eigentlich sehr zufrieden

Im Zuge eines Produkttestes habe ich zwei Packungen Pure Protection umsonst erhalten. Ich finde Sie sehr weich und angenehm. Auch finde ich, dass sie sehr viel Urin speichern kann, ohne, dass der Body oder der Babypoo nass ist. Anders sieht es jedoch aus, wenn der Junge Mann sein Geschäftchen macht. Hier hatten wir schon mehrmals das Problem, dass die Windel nicht gehalten hat. Sie hat die Flüssigkeit nicht aufgesaugt und der Babypoo war mehr als nur nass. Auch lief alles aus und die Kleidung musste gewechselt werden. Ansonsten sind wir sehr zufrieden mit der neuen Pampers. Auch optisch finde ich sie gut gelungen, mal ein anderes Design.

Ursprünglich erschienen auf pampers.de

Abbildung 4: Anzeige Bewertungen bei dm⁵³

Die beiden angezeigten Bewertungen wurden in diesem Fall nicht beim Online-Händler selbst sondern extern auf Pampers.de abgegeben, worüber der Nutzer mit einem entsprechenden Banner unterhalb der Bewertung informiert wird.

Im Rahmen der Sektoruntersuchung hat das Bundeskartellamt die Portale insbesondere dazu befragt, welche Bewertungen in welcher Form dargestellt werden und wie die angezeigten Zusammenfassungen der Bewertungen erstellt werden.

D.3.1. Einzeldarstellung

Die Einzeldarstellung der Bewertungen erfolgt bezüglich des Umgangs mit externen Bewertungen, des Alters von Bewertungen und der Reihenfolge der Anzeige bei allen befragten Portalen bis auf eine Ausnahme (s. u.) in ähnlicher Weise.

Gut ein Drittel aller Portale zeigt neben den eigenen, also über das Portal selbst erhobenen Bewertungen, auch Bewertungen anderer Websites an. In einigen Fällen (bei knapp 20 % der Befragten) handelt es sich dabei schlicht um die Anzeige von Bewertungen, die für das gleiche Portal mit einer anderen Länderdomain

⁵³ <https://www.dm.de/pampers-windeln-pure-protection-groesse-3-midi-6-10-kg-monatsbox-p8001841115917.html> (abgerufen am 01.10.2020).

abgegeben wurden. So hat z. B. OBI angegeben, auf Obi.de auch Bewertungen anzuzeigen, die auf Obi.at abgegeben wurden. Auf Amazon.de werden teilweise Bewertungen angezeigt, die auf Amazon.com abgegeben wurden. Teilweise stammen die Bewertungen auch von anderen konzernverbundenen Portalen, so werden z. B. die Bewertungen von 123Recht.de und Frag-einen-anwalt.de jeweils wechselseitig dargestellt. Einige Portale kennzeichnen Bewertungen, die auf einem anderen, konzerneigenen Portal abgegeben wurden, bei anderen erfolgt keine Kennzeichnung.

Bei knapp 30 % der Befragten erfolgt neben der Darstellung interner auch die Darstellung konzernextern erhobener Bewertungen. Dies basiert in der Regel auf einer vertraglichen Vereinbarung. So bietet der Dienstleister Bazaarvoice seinen Kunden die sog. Syndizierung (Zusammenführung) von Bewertungen an, wodurch Online-Handelsportale, die Bazaarvoice nutzen, die Möglichkeit haben, auch auf der Seite des Herstellers abgegebene Bewertungen für dessen Produkte anzuzeigen, wenn dieser ebenfalls Kunde von Bazaarvoice ist. Beispielsweise werden auf dem Onlineportal Conrad.de für Smart TVs auch Bewertungen angezeigt, die auf der Seite des Herstellers abgegeben wurden. Gekennzeichnet werden die entsprechenden Bewertungen mit dem Hinweis „ursprünglich erschienen auf...“. Eine erneute Filterung findet im Rahmen der Syndizierung nicht statt.

Google und Google Shopping sind von den Befragten diejenigen Portale, die die meisten Kooperationspartner haben, von denen sie Bewertungen übernehmen und anzeigen. Weltweit sind dies etwa 30 Unternehmen, darunter Bazaarvoice, Trusted Shops und Trustpilot, die ihrerseits wiederum mit einer Vielzahl von Unternehmen zusammenarbeiten.⁵⁴ Dabei durchlaufen die angezeigten Bewertungen nach Angaben des Unternehmens sowohl die Filtermechanismen von Google als auch die des Portals, auf dem die Bewertung ursprünglich abgegeben wurde. Externe Bewertungen sind als solche gekennzeichnet, und Nutzer werden in der Regel, wenn sie eine entsprechende Bewertung lesen möchten, auf das jeweilige Portal weitergeleitet.

Die Betreiber der App-Stores haben angegeben, keine externen Bewertungen anzuzeigen.

Auch bereits vor mehreren Jahren abgegebene Bewertungen sind auf den meisten Portalen noch zu sehen. Im Bereich Produktbewertungen haben alle Portale angegeben, Bewertungen anzuzeigen, solange das entsprechende Produkt verfügbar ist. App-Bewertungen werden angezeigt, solange die App verfügbar ist. Nur im Bereich Unternehmensbewertungen haben einige Portale berichtet, dass sie Bewertungen teilweise nach einem gewissen Zeitraum von ein bis mehreren Jahren nicht mehr darstellen. Mithilfe des immer mit der Bewertung angegebenen Datums oder eines anderen Hinweises auf den Zeitpunkt der Abgabe ist das Alter einer Bewertung allerdings in allen drei Bereichen für den Verbraucher transparent, sodass dieser ggfs. für sich entscheiden kann, dass eine Bewertung aufgrund ihres Alters für ihn nicht mehr relevant ist.

Das Alter einer Bewertung ist häufig (mit)entscheidend dafür, wie prominent diese platziert wird: Etwa die Hälfte aller Portale zeigt Bewertungen im Erstranking chronologisch sortiert an („Neueste zuerst“). Die andere Hälfte der Portale inklusive der App-Stores nimmt eine Reihung nach Relevanz vor, wobei auch hier das Alter einer Bewertung immer ein oder sogar das wichtigste Relevanzkriterium ist. Außerdem spielt für die

⁵⁴ <https://support.google.com/google-ads/answer/2375474?hl=de> (abgerufen am 01.10.2020).

Relevanz einer Bewertung regelmäßig eine Rolle, ob die Bewertung von anderen Nutzern als hilfreich erachtet wird (falls dies erhoben wird), die Länge der Bewertung, ob sie ein Foto enthält u. a. Manche Portale platzieren bewusst sowohl eine positive als auch eine negative Bewertung nebeneinander. Viele Portale bieten den Verbrauchern die Möglichkeit, die Reihenfolge der dargestellten Bewertungen nach bestimmten Kriterien zu verändern („Hilfreichste zuerst“, „Schlechteste zuerst“) oder sich nur einen Teil der Bewertungen anzeigen zu lassen (z. B. im Reisebereich: „Bewertungen von Familien“, „Bewertungen von Paaren“). Bei Yelp werden zunächst nur die vom Portal selbst empfohlenen Bewertungen angezeigt. Um die derzeit nicht empfohlenen Beiträge zu sehen, muss ein entsprechender Hinweis angeklickt werden.

Einen Sonderfall stellt das Portal Pearl.de dar. Pearl ist ein Versandhaus für Elektronik- und Lifestyleprodukte, das seine Waren vor allem online über das Portal Pearl.de und einen Shoppingkanal, Pearl.tv, vermarktet. Pearl zeigt auf seiner Seite für einige Produkte „Kundenmeinungen“ bzw. „Kundenbeiträge“ an, wie beispielhaft in nachfolgendem Screenshot dargestellt. Nach eigener Auffassung betreibt das Portal jedoch kein Nutzerbewertungssystem im Sinne der Sektoruntersuchung bzw. zeigt keine Produktbewertungen nach der Definition des Bundeskartellamts⁵⁵ an, denn nach den Angaben des Portals können Kunden über das Portal direkt keine Bewertungen abgeben und es werden auch keine Bewertungen von anderen Seiten dargestellt. Die nähere Betrachtung der Website bestätigt diese Angaben. Die „Kundenmeinungen“ bzw. „Kundenbeiträge“ stammen laut Angaben von Pearl aus Reaktionen auf Qualitätsbefragungen der Kunden im Anschluss an die Versendung online bestellter Produkte. Eine Überprüfung der Kundenmeinungen finde nicht statt.

⁵⁵ Die vom Bundeskartellamt im Auskunftsbeschluss verwendete Definition einer Produktbewertung lautet: „Private Meinungsäußerung zu einem Produkt, die aus einer Sternebewertung und/oder Rezension (Text) besteht, auf einem Internetportal dargestellt wird und im Zusammenhang mit einer Vermittlungs- oder Erwerbsmöglichkeit für ein Produkt steht.“

The screenshot shows the Pearl.de website interface. At the top, there is a navigation bar with the Pearl logo, a search bar, and a shopping cart icon. The main content area is divided into several sections. On the left, there is a navigation menu with categories like 'Dekoration & Wohnaccessoires', 'Sport, Freizeit, Reise, Spielzeug, Games', etc. The central part of the page features a product listing for 'Rosenstein & Söhne Standgrill, Elektro Kugel Grills'. Below the product name, there is a list of features and a detailed product description. To the right of the product description, there is a section for customer reviews, which is circled in red. This section displays '11 Kundenmeinungen für Elektro-Kugel-Grills' and shows three individual reviews with their authors' names and locations. Below the reviews, there is a star rating of 5 stars. To the right of the reviews, there are promotional banners for shipping costs and express delivery.

Abbildung 5: Anzeige Kundenmeinungen/-beiträge auf Pearl.de⁵⁶

D.3.2. Zusammenfassende Darstellung

Für den Nutzer mindestens ebenso relevant wie die einzelnen Bewertungen ist die zusammenfassende Darstellung der Sternebewertung, allein weil diese sehr viel früher angezeigt und daher auch ohne explizites Anklicken vom Nutzer wahrgenommen wird. Marktteilnehmer haben darauf hingewiesen, dass die Zusammenfassung auch das Verhalten von Nutzern beeinflusst, die sich gar nicht für Bewertungen interessieren und sich einzelne Bewertungen daher auch nicht anschauen. Dadurch, dass die Zusammenfassung der Bewertungen in Form farbiger Sterne direkt bei der Produktdarstellung angezeigt werde, könne man sich dem Einfluss dieser Zusammenfassung gar nicht entziehen.

Entsprechend zeigen alle befragten Portale eine Bewertungszusammenfassung an, in der Regel einen Durchschnitt der einzelnen Sternebewertungen wie in der obigen Abbildung dargestellt. Der angegebene Durchschnitt ist bis auf wenige Ausnahmen ein einfacher arithmetischer Mittelwert, in den alle Bewertungen mit dem gleichen Gewicht einfließen. Bei einigen Portalen werden ältere Bewertungen nicht mit eingezogen, die Grenze liegt dabei zwischen sechs Monaten und vier Jahren. Die befragten Portale, die die Möglichkeit der Syndizierung von Bewertungen über Bazaarvoice nutzen, beziehen die darüber zusätzlich dargestellten externen Bewertungen auch in die zusammenfassende Darstellung der Bewertungen mit ein.

Das Portal Yelp bezieht nur solche Bewertungen in die Zusammenfassung ein, die aufgrund der eigenen Richtlinien und einer internen Überprüfung als „empfohlen“ eingestuft werden; die übrigen Bewertungen können

⁵⁶ <https://www.pearl.de/a-NX5150-3510.shtml?query=kugelgrill> (abgerufen am 01.10.2020).

jedoch zusätzlich abgerufen werden. Das Vorgehen von Yelp kann dazu führen, dass nur wenige, im Extremfall nur eine einzige Bewertung eines Anbieters als „empfohlen“ eingestuft wird. Wenn es sich dabei um eine negative Bewertung handelt, ist dementsprechend auch der dem Nutzer plakativ am Anfang der Seite ausgewiesene Bewertungsdurchschnitt negativ. Selbst wenn der Anbieter parallel dazu eine ganze Reihe von 5-Sterne-Bewertungen aufweisen kann, die aber allesamt als „nicht empfohlen“ eingestuft wurden, wären diese nur in der weniger sichtbaren Kategorie zu finden. Vergleichbare Fallkonstellationen waren in der Vergangenheit Gegenstand verschiedener gerichtlicher Auseinandersetzungen. Im Januar 2020 bestätigte der Bundesgerichtshof die Rechtmäßigkeit der Vorgehensweise von Yelp.⁵⁷

In die Berechnung der Durchschnittsbewertungen, die auf Google oder Google Shopping als Teil des Suchergebnisses angezeigt werden, sind externe Bewertungen für Produkte und Händler mit einbezogen, externe Bewertungen für Hotels jedoch nicht. Bei Händlern, für die sowohl gewöhnliche Nutzerbewertungen als auch Bewertungen über das Test Shopping Program von Google (vgl. dazu Kapitel D.1.4.) vorliegen, gehen die Bewertungen durch das Test Shopping Program in die Durchschnittsbewertung ein. Damit für Verkäufer überhaupt eine Durchschnittsbewertung als Teil des Suchergebnisses angezeigt wird, müssen zudem insgesamt mindestens 100 Bewertungen aus den letzten 12 Monaten vorliegen und bei den allgemeinen Suchergebnissen – nicht auf Google Shopping – muss die Durchschnittsbewertung des Verkäufers mindestens 3,5 Sterne betragen.⁵⁸

Eine Ausnahme bei der Durchschnittsberechnung bildet das Portal Trustpilot, dessen zusammenfassende Darstellung aller Bewertungen, der „TrustScore“, als Bayes-Durchschnitt berechnet wird. Der Bayes-Durchschnitt ist ein arithmetischer Mittelwert, in den neben den tatsächlichen Beobachtungen noch ein a priori festgesetzter Mittelwert mit einbezogen wird. So würde z. B. bei der Ermittlung des Durchschnittsgewichts aller Männer einer deutschen Volleyballmannschaft neben dem Gewicht von einem oder zweier seiner Mitglieder auch noch das Durchschnittsgewicht aller deutschen Männer mit berücksichtigt. Ein Bayes-Durchschnitt kann ein geeignetes Maß für die Schätzung eines Mittelwertes darstellen, wenn nur sehr wenige Beobachtungen vorliegen, aber bekannt ist, wie groß der Mittelwert der Beobachtungen ungefähr sein könnte. In dem genannten Beispiel vermeidet man, dass falsche Rückschlüsse von einem Einzelfall auf die ganze Mannschaft gezogen werden, was dann sinnvoll erscheint, wenn unglücklicherweise der schwerste oder der leichteste aus der Mannschaft als einzige Beobachtung zur Verfügung steht.⁵⁹ Bei Trustpilot ergibt sich für den Fall, dass ein Unternehmen erst *eine* Bewertung bekommen hat, die einen Stern aufweist, aufgrund der Berücksichtigung des Mittelwerts von üblicherweise etwas mehr als 3 Sternen für Unternehmensbewertungen ein TrustScore für das Unternehmen von 3,2. In die Berechnung des Durchschnitts fließt außerdem mit ein, wie alt die Bewertungen für ein Unternehmen sind und wie häufig das Unternehmen Bewertungen erhält.⁶⁰ Die Berechnungsmethode für die zusammenfassende Darstellung wird auf der Seite von Trustpilot

⁵⁷ BGH, Urt. v. 14.01.2020, Az. VI ZR 496/18.

⁵⁸ <https://support.google.com/google-ads/answer/2375474?hl=de> (abgerufen am 01.10.2020).

⁵⁹ https://en.wikipedia.org/wiki/Bayesian_average (abgerufen am 01.10.2020).

⁶⁰ <https://support.trustpilot.com/hc/en-us/articles/201748946-TrustScore-explained-How-is-the-TrustScore-calculated-> (abgerufen am 01.10.2020).

dargestellt, erläutert und begründet. Die Methode des Bayes-Durchschnitts wird auch von Filmstarts.de verwendet, das bei der ersten Bewertung 20 anonyme 3-Sterne-Bewertungen hinzufügt. Das Ziel dieses Vorgehens sei es, extreme Bewertungen zu relativieren und Manipulationen zu vermeiden. Zwei Portale gaben an, aus diesem Grund eine zusammenfassende Bewertung erst bei einer Mindestanzahl an Bewertungen anzuzeigen. Ähnlich zeigt das interne Ablaufschema eines weiteren Portals, dass negative Bewertungen zunächst zurückgestellt, also nicht veröffentlicht werden, bis das betreffende Produkt eine bestimmte Mindestanzahl von Bewertungen erhalten hat. Hier erfolgt die Zurückstellung demnach asymmetrisch, sie diskriminiert negative Bewertungen, wohingegen positive Bewertungen keine Gefahr laufen, als Ausreißer zurückgestellt zu werden. Auch auf diesem Wege kann der Einfluss von einzelnen Extrembewertungen relativiert werden.

Das Portal Pearl.de zeigt auf den Seiten, auf denen „Kundenmeinungen“ bzw. „Kundenbeiträge“ dargestellt werden, auch eine zusammenfassende Darstellung an, wie der folgende beispielhafte Screenshot illustriert. Es handelt sich um einen etwas weiter unten liegenden Ausschnitt der gleichen Seite wie in Abbildung 6. Pearl selbst hat allerdings angegeben, auf seiner Seite keine durchschnittliche Bewertung anzuzeigen.

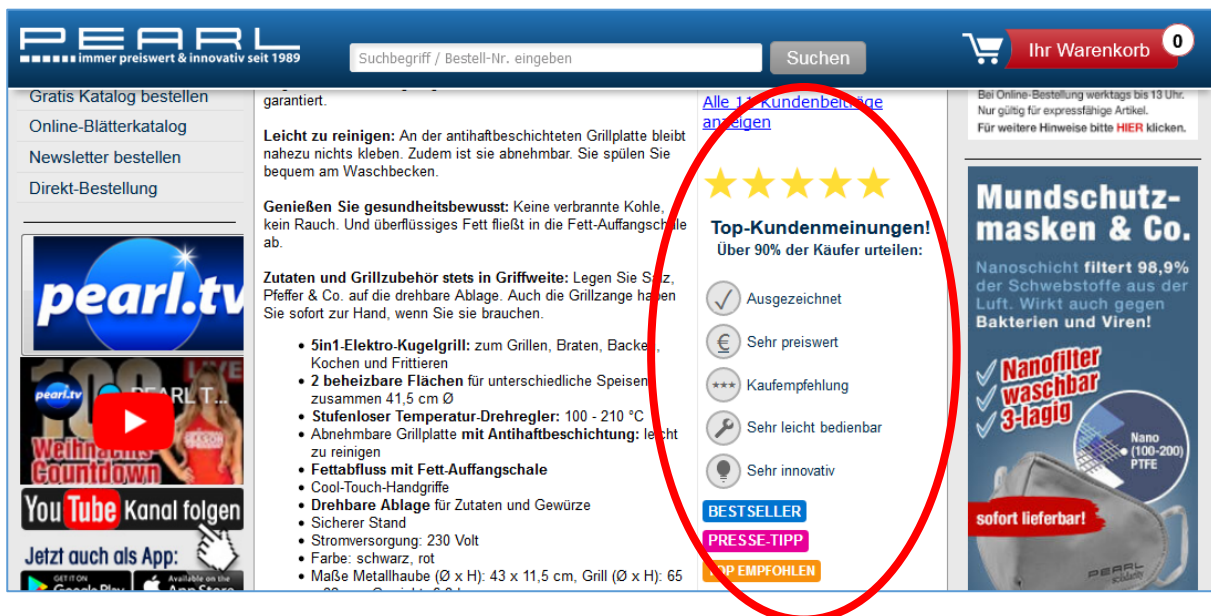


Abbildung 6: Anzeige Zusammenfassung Kundenmeinungen auf Pearl.de⁶¹ (Quelle: Pearl, 2020)

D.3.3. Wesentliche Erkenntnisse

Hinsichtlich der Einzeldarstellung der Bewertungen haben die Ermittlungen im Rahmen der Sektoruntersuchung bis auf eine Ausnahme (s. F.6.1.) keine Auffälligkeiten ergeben. Die weit überwiegende Mehrheit der befragten Portale zeigen sämtliche (gefilterten) Bewertungen an, kennzeichnen Bewertungen, die von anderen Websites übernommen wurden, sortieren die Bewertungen nach Relevanz bzw. Aktualität und bieten dem Verbraucher eigene Sortier- oder Filtermöglichkeiten für die angezeigten Bewertungen an.

⁶¹ <https://www.pearl.de/a-NX5150-3510.shtml?query=kugelgrill> (abgerufen am 01.10.2020).

Der zusammenfassenden Darstellung der Sternebewertungen kommt aufgrund des starken Einflusses auf den Verbraucher eine besondere Bedeutung zu. Problematisch können insbesondere solche Zusammenfassungen sein, bei denen nur ein Teil der Bewertungen einbezogen wird, die den Filterprozess erfolgreich durchlaufen haben oder bei denen künstliche Bewertungen hinzugefügt werden. Kritisch zu beurteilen sind darüber hinaus zusammenfassende Darstellungen, die nicht auf dem arithmetischen Mittelwert der Bewertungen beruhen, sondern unterschiedliche, ggfs. verzerrende Gewichtungen vornehmen. Derartige Beeinflussungen der zusammenfassenden Bewertung durch das Portal müssen für den Verbraucher klar gekennzeichnet sein und sollten ggfs. durch den Verbraucher ausgeschaltet werden können. Dies gilt auch für vom Portal mittels eigener Tests erstellter Bewertungen. Um extreme Durchschnittsbewertungen aufgrund weniger Einzelbewertungen zu vermeiden, können Portale – wie von einigen bereits praktiziert – die Darstellung einer Zusammenfassung vom Vorliegen einer Mindestzahl an Bewertungen abhängig machen und dies ebenfalls gegenüber dem Verbraucher kenntlich machen (s. Kapitel F.6.2.).

E. Interessenlagen und Verhaltensweisen der Marktteilnehmer

Nutzerbewertungen sind in erster Linie Meinungsäußerungen von Verbrauchern für Verbraucher im Internet. Durch den Einfluss, den die Bewertungen auf die Entscheidung der Verbraucher haben, sind sie jedoch auch für die Anbieter von Bedeutung, die entweder selbst online Produkte anbieten oder auf einem Bewertungsportal bewertet werden. Der starke Einfluss der Nutzerbewertungen auf das Kaufverhalten der Verbraucher führt dazu, dass die Anbieter teilweise versuchen, diesen Zugang zum Verbraucher bewusst zu gestalten und zu beeinflussen. Falls die Anbieter auf Plattformen bewertet werden, stehen letztere daher vor der Herausforderung, rechtskonforme Spielregeln für diese Gestaltung einzuführen und deren Einhaltung zu überwachen. Die Zusammenhänge zwischen den verschiedenen Beteiligten illustriert auch die folgende Abbildung: Zahlreiche Verbraucher lesen auf Onlineportalen und -plattformen Bewertungen und informieren sich damit über bestimmte Leistungen, die dann von den Anbietern online oder offline erbracht werden. Ein kleiner Teil der Verbraucher verfasst auch selbst Bewertungen und stellt diese auf einem Onlineportal ein. Dort stehen sie dann wieder allen anderen Verbrauchern bzw. Nutzern des jeweiligen Portals zur Verfügung.

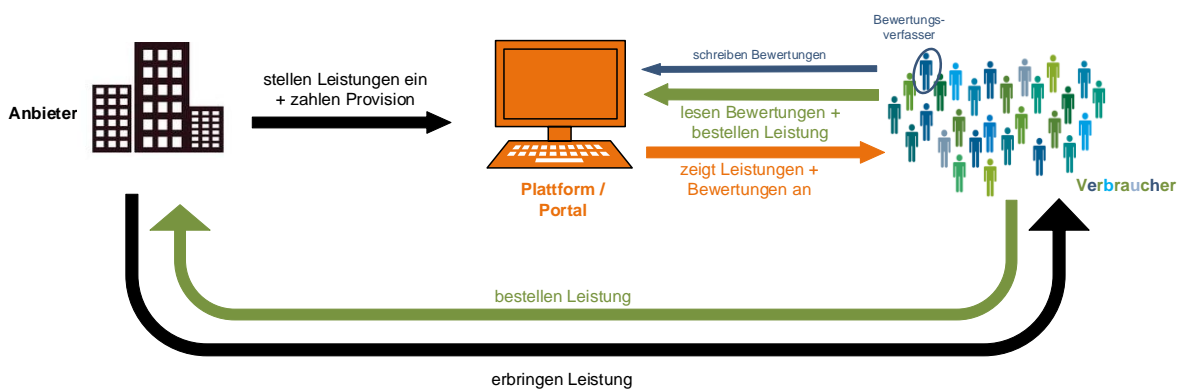


Abbildung 7: Marktteilnehmer bei Nutzerbewertungen

Das folgende Kapitel beleuchtet die Rolle der einzelnen Marktteilnehmer und ihre Beziehungen untereinander. Neben Verbrauchern, Anbietern und Plattformen bzw. Portalen werden dabei als zusätzliche Akteure sog. Bewertungsvermittler vorgestellt, die im Auftrag von Anbietern zusätzliche Nutzer akquirieren, die entgeltlich als „Produkttester“ Bewertungen schreiben.

E.1. Verbraucher

Nutzerbewertungen sind im Kern eine Art der Kommunikation zwischen Verbrauchern, die sich mithilfe von Bewertungen, die sie selbst verfassen und lesen, über Produkte und Anbieter austauschen. Diese Form der Kommunikation ergänzt und ersetzt u. a. den mündlichen Austausch mit Freunden und Bekannten über Erfahrungen mit Produkten und Dienstleistungen und stellt für Verbraucher gerade in der digitalen Welt eine wichtige Möglichkeit dar, Informationen darüber zu bekommen.

Im Unterschied zu Werbung steht hinter Nutzerbewertungen in der Regel kein wirtschaftlicher Anreiz: Werbung dient – auch wenn sie informativ ist – letztlich der Umsatzsteigerung; Bewertungen sind diesbezüglich im Ausgangspunkt neutral. Von dritter Seite durchgeführte, professionelle Produkttests wie z. B. von Stiftung Warentest sind grundsätzlich auch neutral, verwenden aber häufig andere Kriterien für ihre Bewertung als Verbraucher, die Nutzerbewertungen verfassen: Bei Stiftung Warentest wird z. B. ein Produkt auf Haltbarkeit und Schadstoffgehalt geprüft, eine Nutzerbewertung beschreibt dagegen, wie intuitiv ein Gerät zu bedienen ist oder ob ein Kleidungsstück gut aussieht.

Das Verhalten der Verbraucher beim Lesen bzw. Schreiben von Nutzerbewertungen war in der Vergangenheit bereits Gegenstand verschiedener Studien und Verbraucherbefragungen im In- und Ausland.⁶² In Deutschland haben sowohl der Verband Bitkom⁶³ als auch die Gesellschaft für Konsumforschung (GfK) bereits mehrmals Befragungen insbesondere zur Bedeutung von Nutzerbewertungen für die Entscheidungsfindung der Verbraucher durchgeführt.⁶⁴ Befragungen durch die Verbraucherzentrale Bayern⁶⁵ und durch Statista⁶⁶, letztere speziell zu Amazon, legten den Schwerpunkt auf das Vertrauen der Verbraucher in Nutzerbewertungen.

Das Bundeskartellamt hat zu diesem Thema bislang keine eigene Verbraucherbefragung durchgeführt. Im Rahmen der vorliegenden Sektoruntersuchung wurden jedoch die ausgewählten Portale um die Übersendung von Unterlagen zur Auswertung des Verbraucherverhaltens in Bezug auf Nutzerbewertungen gebeten. Insbesondere zu Unternehmensbewertungen im Reisebereich, aber auch zu Bewertungen lokaler Dienstleister haben einzelne Portale in diesem Zusammenhang Ergebnisse eigener Befragungen vorgelegt, die sowohl das Verhalten der Verbraucher als Leser von Nutzerbewertungen wie auch das Verhalten der Verbraucher als Verfasser von Nutzerbewertungen betreffen.⁶⁷ Teilweise zielten die von den Portalen vorgelegten Auswertungen des Verbraucherverhaltens auf die Verbesserung der eigenen Leistung ab, z. B. durch Rückschlüsse auf die Qualität bestimmter Produkte oder Serviceleistungen oder durch die Optimierung des eigenen Bewertungssystems für unterschiedliche Nutzertypen. Teilweise betrafen die Auswertungen allerdings auch allgemeine Erkenntnisse zum Verhalten der Verbraucher in Bezug auf Nutzerbewertungen. Da sich die

⁶² Als Übersicht zu zahlreichen ausländischen Studien vgl. etwa: <https://learn.g2.com/customer-reviews-statistics> (abgerufen am 01.10.2020).

⁶³ Bitkom Research, Kundenbewertungen sind wichtigste Kaufhilfe, 2017 und <https://www.bitkom.org/Presse/Presseinformation/Jeder-Zweite-liest-Online-Bewertungen-vor-dem-Kauf> (abgerufen am 01.10.2020).

⁶⁴ GfK-Umfrage von Greven Medien, 2017 und 2019.

⁶⁵ Verbraucherzentrale Bayern, Online-Bewertungen aus Sicht von Verbrauchern, 2017, <https://www.marktwaechter.de/nur-wenige-verbraucher-verlassen-sich-voll-und-ganz-auf-online-bewertungen> (abgerufen am 01.10.2020).

⁶⁶ Pepper Media Holding, Vertrauen Sie Bewertungen auf Amazon? Statista, <https://de.statista.com/statistik/daten/studie/1057375/umfrage/vertrauen-in-amazon-bewertungen-in-deutschland/> (abgerufen am 01.10.2020).

⁶⁷ Holidaycheck/Tomorrow Focus AG, Die Psychologie des Bewertens, 2014 und 2016.

TripAdvisor und Ipsos MORI, The Power of Reviews, 2019.

Das Örtliche Servicegesellschaft, Umfrage Mente>Factum, Q4/2019.

vorgelegten Unterlagen zum Teil auf unterschiedliche Branchen und/oder Länder bezogen, sind deren Ergebnisse nicht immer vergleichbar. Die daraus resultierenden Erkenntnisse lassen jedoch in einigen Punkten allgemeine Schlussfolgerungen zum Verbraucherverhalten zu.

Neben der Anforderung von Unterlagen wurden die ausgewählten Portale sowie weitere Marktteilnehmer im Rahmen der Sektoruntersuchung auch unmittelbar zu ihren Erkenntnissen in Bezug auf das Verbraucherverhalten bei Nutzerbewertungen befragt. Die Fragen bezogen sich darauf, in welchen Situationen bzw. von welchen Verbrauchern überhaupt Nutzerbewertungen abgegeben werden und darauf, wie häufig welche Sternebewertungen vergeben werden.

E.1.1. Verbraucher als Leser von Nutzerbewertungen

Die verschiedenen Studien zeigen übereinstimmend, dass Nutzerbewertungen im Internet von einem Großteil der Verbraucher gelesen werden und die Verbraucher auch in ihrer Entscheidung beeinflussen. Beispielfähig gaben in der letzten Bitkom-Befragung vom Oktober 2019 insgesamt 56 % der Befragten an, Bewertungen als Entscheidungshilfe zu verwenden. Allerdings äußerten sich in Befragungen der Verbraucherzentrale Bayern und von Statista die Befragten durchaus auch kritisch: 12 % der Internetnutzer verlassen sich demnach gar nicht oder eher nicht auf Bewertungen, speziell Bewertungen auf Amazon misstrauen nach eigenen Angaben sogar etwa ein Viertel aller Nutzer. Die Hauptmotivationsfaktoren der Verbraucher sind lt. verschiedenen Studien die Orientierung, die Vertrauensbildung und die Verbesserung bzw. Absicherung der Kaufentscheidung. Im internationalen Vergleich scheint dabei das Vertrauen deutscher Verbraucher in Nutzerbewertungen insgesamt etwas geringer zu sein als das Vertrauen der Verbraucher in anderen Ländern, insbesondere den USA.

Jüngere Verbraucher und Frauen nutzen Bewertungen offenbar häufiger als Ältere und Männer. Jüngere Verbraucher schätzen dabei auch in stärkerem Maße ergänzende Fotos und Videos. Die Auswertungen legen zudem nahe, dass Verbraucher insbesondere aktuelle Bewertungen lesen und Extrembewertungen teilweise ignorieren. Auch die persönlichen Antworten von Anbietern auf Bewertungen spielen für die Vertrauensbildung der Verbraucher eine große Rolle. Angebote, die (noch) keine Bewertung erhalten haben, werden nach den Studien vom Verbraucher seltener ausgewählt als Angebote, die vereinzelte negative Bewertungen erhalten haben. Eine wichtige Rolle für die Kaufentscheidung des Verbrauchers spielen Bewertungen vor allem bei „unbekannten“ Anbietern wie unabhängigen Restaurants oder Hotels, die keiner Kette angehören.

E.1.2. Verbraucher als Verfasser von Nutzerbewertungen

Verschiedene Untersuchungen deuten darauf hin, dass der Anteil an Nutzern, der überhaupt eine Bewertung abgibt, je nach Situation zwar unterschiedlich hoch ist, in der Regel aber trotz aller Anreize eher im niedrigen

einstelligen Prozentbereich liegt.⁶⁸ Einer der Marktteilnehmer sprach dazu von einer „1%-Regel“ für aktive Nutzer, d. h. ca. 1 % aller Nutzer gibt tatsächlich Bewertungen ab. Die ausgewerteten Studien zeigen auch, dass die meisten Bewertungen für Hotels abgegeben werden, gefolgt von Restaurants, Händlern, Handwerkern und Ärzten. Produktbewertungen werden vor allem für Elektronikartikel und Kleidung abgegeben. Auch nach den Ergebnissen der Sektoruntersuchung Vergleichsportale⁶⁹ des Bundeskartellamtes sind die Anteile der bewertenden Nutzer je nach Branche unterschiedlich hoch: So geben danach etwa 5 % der Kunden auf Energievergleichsportalen eine Bewertung ab, aber 14,5 % der Kunden auf Hotelvergleichsportalen. Die Motivation der Bewertungsverfasser besteht gemäß den Studien zumeist darin, anderen Nutzern zu helfen, die eigenen Emotionen (Dank, Verärgerung) zum Ausdruck zu bringen oder dem Anbieter ein Feedback zu geben. Im Einklang damit gibt es Hinweise, dass die Bereitschaft, kostenlose Produkte zu bewerten, höher ist, weil Verbraucher sich ggfs. verpflichtet fühlen, sich dem Anbieter gegenüber erkenntlich zu zeigen. So werden in App-Stores für kostenlose Apps relativ mehr Bewertungen abgegeben als für kostenpflichtige Apps.⁷⁰ Der Hauptgrund für das Nicht-Verfassen einer Bewertung liegt nach den Erkenntnissen aus verschiedenen Studien in allen Bereichen klar in dem damit verbundenen Aufwand.

Wenn die wenigen Nutzer, die Bewertungen abgeben, eine Leistung systematisch anders bewerten als die große Mehrheit der Nutzer, die keine Bewertung abgibt, sind die abgegebenen Bewertungen verzerrt und entsprechen nicht der typischen Nutzermeinung. Wenn den Lesern von Nutzerbewertungen diese Verzerrung nicht bewusst ist, kann sie im Ergebnis zu falschen Verbraucherentscheidungen führen. In einer empirischen Untersuchung wurde dies bereits für die Bewertung von Büchern auf Amazon.com gezeigt:⁷¹ Diejenigen, die ein neu erschienenes Buch sofort kaufen und bewerten, geben eine überdurchschnittlich gute Bewertung ab, vielleicht weil sie ohnehin den Autor des Buches besonders schätzen. Dieser Selektionseffekt wird von den Kunden, die das Buch etwas später kaufen, nicht hinreichend berücksichtigt, sie kaufen das Buch aufgrund der (zu) positiven Bewertungen und sind dann überwiegend enttäuscht. Im Zeitablauf erhält ein neu erschienenes Buch so zunächst besonders positive und dann besonders negative Bewertungen, bevor die Bewertungen sich auf einen Mittelwert einpendeln. Entscheidend ist, dass die Käufer auf Amazon.com zumindest für den untersuchten Datensatz nicht in der Lage waren, diesen Effekt zu erkennen und bei ihrer Entscheidung zu berücksichtigen; sie wurden aufgrund der überwiegend positiven Erstbewertungen dazu gebracht, ein Buch zu kaufen, das sie dann zumindest teilweise enttäuschte.

⁶⁸ Die o. g. Studie der Verbraucherzentrale Bayern aus dem Jahr 2017 hatte allerdings ermittelt, dass sieben von zehn Verbrauchern immer, häufig oder manchmal online Bewertungen abgeben.

⁶⁹ *Bundeskartellamt*, Sektoruntersuchung Vergleichsportale, S. 119 ff., April 2019, https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Sektoruntersuchungen/Sektoruntersuchung_Vergleichsportale_Bericht.html (abgerufen am 01.10.2020).

⁷⁰ *Bond, He, Wen*, Speaking for "Free": Word of Mouth in Free- and Paid-Product Settings, *Journal of Marketing Research*, Band 56, Ausgabe 2, S. 276-290, 2019.

⁷¹ *Li, Hitt*, Self Selection and Information Role of Online Product Reviews, *Information Systems Research*, Band 19, Nr.4, S. 459-474, Dezember 2008.

Im Rahmen der Sektoruntersuchung wurden die einbezogenen deutschsprachigen Portale und sonstige Marktteilnehmer direkt gefragt, welche Nutzer typischerweise Bewertungen abgeben bzw. in welchen Situationen eher Bewertungen abgegeben werden. Dazu wurden von den Befragten verschiedene Aspekte genannt. Grundsätzlich würden Angebote eher bewertet, wenn die Leistung für den Nutzer mit starken Emotionen verbunden ist. Das betrifft z. B. besonders teure Anschaffungen, Prestigeprodukte, Babyartikel oder bestimmte Bücher. Tendenziell würden mehr Bewertungen abgegeben, wenn die Nutzer Zeit haben. Ein Onlineshop berichtete, dass sonntags die meisten und dienstags die wenigsten Bewertungen abgegeben würden. Ein Reiseportal gab an, dass die meisten Bewertung von Reisenden aus der Kategorie „Familie“ abgegeben würde, gefolgt von Paaren und Gruppen. Die wenigsten Bewertungen würden von Alleinreisenden, vermutlich überwiegend Geschäftsreisenden, abgegeben. Würden Nutzer per Mail aufgefordert, Bewertungen abzugeben, sei die Antwortquote bei Frauen höher als bei Männern, bei Älteren höher als bei Jüngeren. Dies hänge möglicherweise damit zusammen, dass sich Frauen und Ältere stärker verpflichtet fühlen, einer entsprechenden Aufforderung nachzukommen. Immer wieder wurde von den Befragten darauf hingewiesen, dass es die Extremerfahrungen sind, die dazu führen, dass Nutzer eine Bewertung abgeben: Bewertet wird, wenn man besonders zufrieden oder besonders unzufrieden ist. Die Mehrheit der Erfahrungen seien zwar keine Extremsituationen, doch finde sich diese Mehrheit in den Nutzerbewertungen nur sehr schwach wieder.

Darüber hinaus wurden die ausgewählten Portale gefragt, wie häufig im Jahr 2018 welche Anzahl an Sternen vergeben wurde. Die folgende Abbildung zeigt die Verteilung der Sterne bei den befragten Portalen für Produkt-, Unternehmens- und App-Bewertungen für diejenigen der befragten Portale, die eine Fünferskala verwenden und eine Verteilung angeben konnten. Die Verteilung bestätigt die Aussagen von Marktteilnehmern bezüglich der Extremsituationen zumindest im Positiven: Zwei Drittel aller angezeigten Bewertungen sind fünf Sterne. Mitarbeiter des international tätigen Bewertungsdienstleister Bazaarvoice haben sogar berichtet, dass nach ihren Erfahrungen der Anteil der 5-Sterne-Bewertungen in Deutschland mit 70 % etwas höher sei als im globalen Durchschnitt, der bei 67 % liege. Im Negativen spiegelt sich die Aussage zu den Extremsituationen dagegen nicht wider: Die zweithäufigste Bewertung ist nicht etwa ein Stern, sondern sind vier Sterne. Erst an dritthäufigster Stelle steht ein Stern, diese Bewertungen machen nur 6 bis 7 % aller angezeigten Bewertungen aus. Insgesamt zeigt die Verteilung der Bewertungen, dass die große Mehrheit der Kunden positiv bewertet. Über alle drei Kategorien hinweg werden entfallen mehr als 80 % der Bewertungen auf vier oder fünf Sterne.

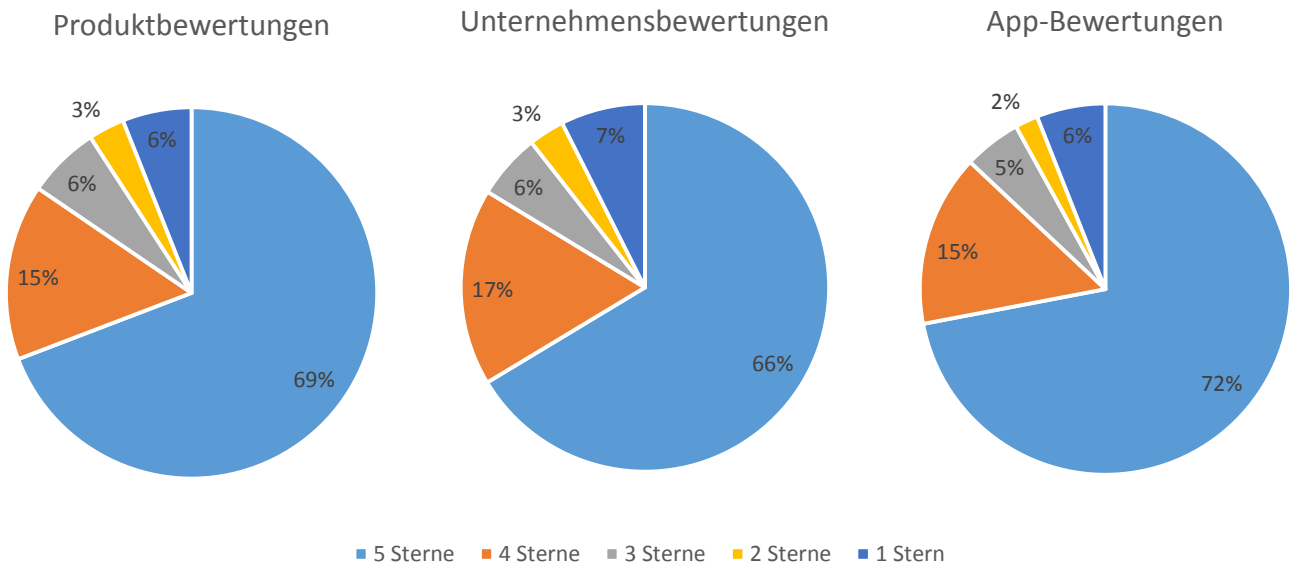


Abbildung 8: Verteilung der Bewertungen auf Sterne in den drei Befragungsgruppen

Zu berücksichtigen ist dabei, dass die angezeigten Bewertungen nur diejenigen sind, die nicht durch Filter entfernt wurden. Es ist davon auszugehen, dass der Anteil negativer Bewertungen an allen abgegebenen Bewertungen etwas höher als dargestellt ist, diese Bewertungen dann aber mit höherer Wahrscheinlichkeit als positive Bewertungen im Filter zurückgehalten werden, z. B. weil sie tendenziell mehr Schimpfworte enthalten oder weil negative Bewertungen kritischer geprüft werden (siehe Kapitel F.5. und F.6.). Aber auch unter Berücksichtigung dieses Effekts wäre immer noch die deutliche Mehrheit aller Bewertungen positiv.

Die folgende Abbildung zeigt den Anteil an 5- und 1-Sterne-Bewertungen bei den befragten Portalen im Bereich Produktbewertungen. Mit einbezogen sind nur diejenigen Portale, die eine Fünferskala zur Bewertung verwenden und eine Verteilung angeben konnten.

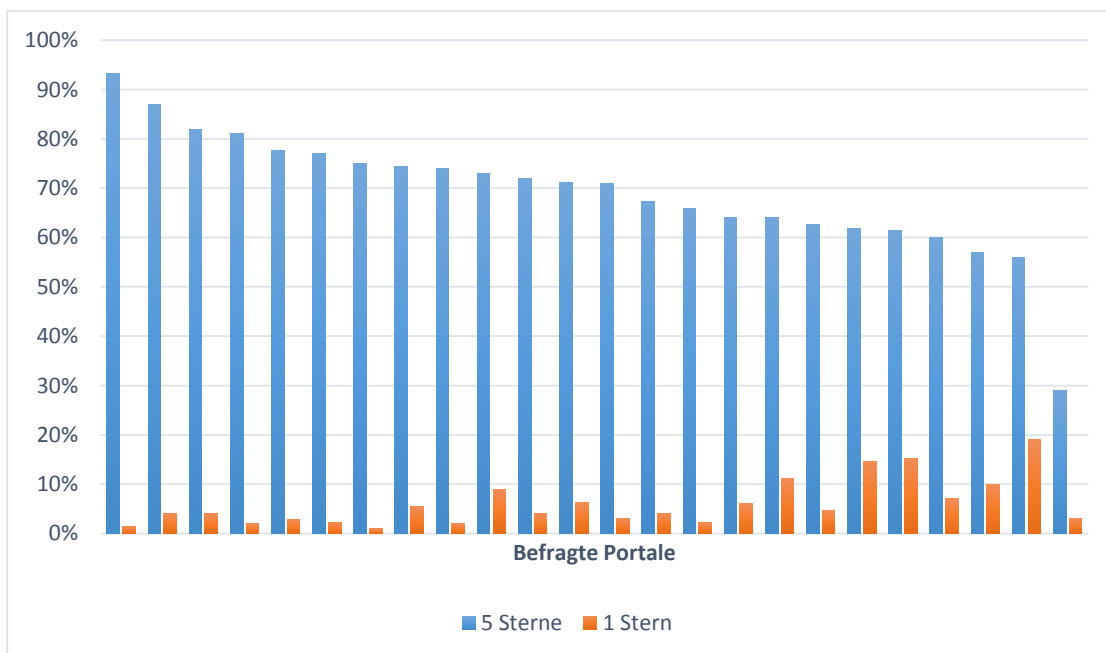


Abbildung 9: Anteil der Produktbewertungen mit fünf Sternen bzw. einem Stern

Der Anteil der 5-Sterne-Bewertungen liegt bei fast allen Portalen über 50 %. Nur bei einem Portal werden in weniger als 30 % aller Fälle fünf Sterne vergeben. Es handelt sich dabei um Maxdome, ein Video-on-Demand-Portal, auf dem Filme bewertet werden können, wobei nur Sterne vergeben werden, aber keine Rezensionstexte. Bei diesem Portal werden am häufigsten 3-Sterne-Bewertungen vergeben, d. h. die Zuschauer bewerten einen Film als „mittelmäßig“.

Mit einer Quote von über 90 % die meisten 5-Sterne-Bewertungen erzielt Etsy, eine Plattform für den Kauf und Verkauf von handgemachten Produkten, Vintage und Künstlerbedarf. Der sehr hohe Anteil an sehr guten Bewertungen hängt hier möglicherweise mit dem Charakter der Plattform zusammen, der durch einen hohen Anteil weiblicher, in Heimarbeit tätiger Anbieter und selbstgemachter Unikate geprägt ist, die bei den Nutzern vermutlich hohe Sympathiewerte genießen. Viele der Shops auf Etsy sind von der Plattform Dawanda übertragen worden, nachdem Dawanda 2018 geschlossen wurde.

Einen höheren Anteil an 5-Sterne-Bewertungen erhalten bei den Befragten auch Portale aus dem Medizinbereich, wohingegen Portale aus dem Baumarktbereich unterdurchschnittlich viele Top-Bewertungen erhalten.

Die folgende Abbildung zeigt den Anteil an 5- und 1-Sterne-Bewertungen bei den befragten Portalen im Bereich Unternehmensbewertungen. Wiederum sind nur diejenigen Portale einbezogen, die eine Fünferskala zur Bewertung verwenden und eine Verteilung angeben konnten.

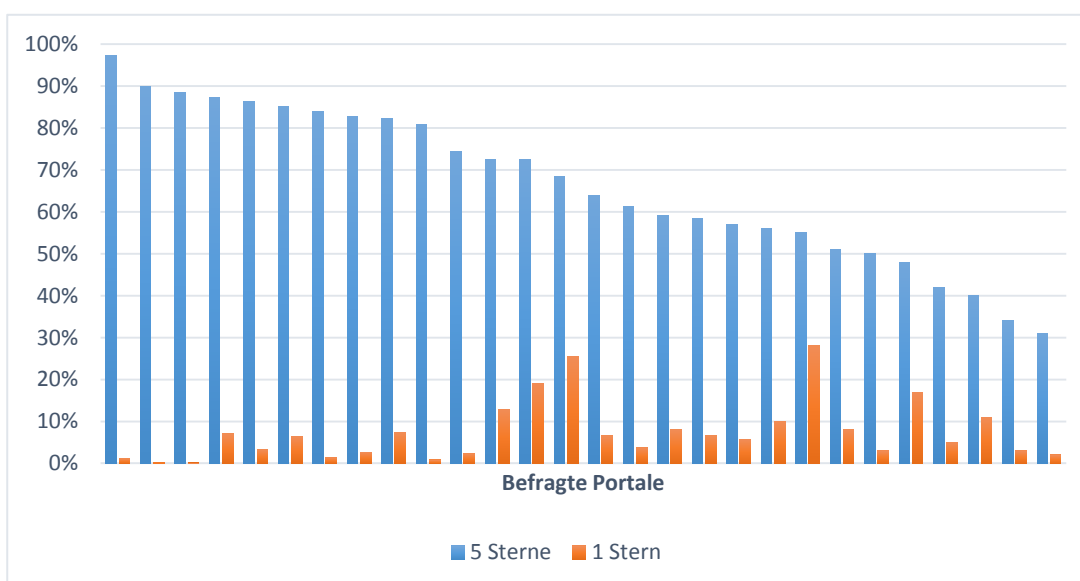


Abbildung 10: Anteil der Unternehmensbewertungen mit fünf Sternen bzw. einem Stern

Auch hier ist das einzige Portal, das mehr als 90 % 5-Sterne-Bewertungen aufweist, eine Plattform mit persönlichem Charakter; es handelt sich um Kleiderkreisel. Kleiderkreisel ist eine Plattform, auf der nur nicht-gewerbliche Verkäufer zugelassen sind und auf der sich Käufer und Verkäufer gegenseitig bewerten können. Einen geringen Anteil an 5-Sterne-Bewertungen weisen in diesem Bereich Plattformen aus dem Reisebereich auf. Auffällig ist bei den betreffenden Portalen ein hoher Anteil an 4-Sterne-Bewertungen. Den höchsten Anteil von 1-Sterne-Bewertungen - über 25 % - weisen Dasoertliche.de und Dastelefonbuch.de auf. Die Aus-

sagekraft der Bewertungen könnte allerdings in diesem Fall höher sein als bei Portalen, auf denen quasi überhaupt keine 1-Sterne-Bewertungen angezeigt werden und fast ausschließlich 5-Sterne-Bewertungen vorliegen.

E.1.3. Wesentliche Erkenntnisse

Insbesondere in der Rechtsprechung, aber auch in der öffentlichen Diskussion um Bewertungen liegt der Fokus oft auf den sog. Fake-Bewertungen und dem damit möglicherweise verbundenen Schaden für den Verbraucher.⁷² Ein Problem mit Bewertungen wird dann gesehen, wenn nicht sichergestellt ist, dass Nutzerbewertungen ausschließlich von anderen Nutzern verfasst wurden, frei und unbeeinflusst vom Anbieter des Produkts, und im Umkehrschluss ist alles in Ordnung, wenn dies sichergestellt ist. Die Ermittlungsergebnisse der Sektoruntersuchung machen jedoch deutlich, dass die mangelnde Repräsentativität der dargestellten Bewertungen ein zweites, mindestens ebenso relevantes Problem darstellen kann:

Es ist unbestritten, dass Nutzerbewertungen im Internet von zahlreichen Verbrauchern gelesen und in die Entscheidungsfindung einbezogen werden. Die ausgewerteten Studien und die Ermittlungen haben jedoch bestätigt, dass nur ein sehr kleiner Teil der Nutzer überhaupt eine Bewertung abgibt und dass das „Ob“ einer Bewertung von der jeweiligen Branche, vom Portal sowie von der individuellen Situation des Nutzers abhängt. Darüber hinaus zeigen die Ermittlungen, dass auch die Anzahl der vergebenen Sterne, also das „Wie“ der Bewertung, von der persönlichen Situation des Bewertungsverfassers sowie den Umständen der Transaktion abhängen, was insbesondere zu dem dargestellten großen Anteil der Extrembewertungen führt.

Je geringer der Anteil der Nutzer, die überhaupt Bewertungen abgeben, und je stärker der Einfluss individueller Umstände auf die Art der Bewertung, desto größer ist die mögliche Verzerrung der dargestellten Bewertungen. Die Portale selbst können mittels ihres Bewertungssystems die mangelnde Repräsentativität der abgegebenen und dargestellten Bewertungen allerdings nur bedingt beeinflussen. Grundsätzlich ist jedoch davon auszugehen, dass solche Bewertungssysteme, bei denen die Betreiber viele und wirkungsvolle Anreize setzen, um Nutzer von der Bewertungsabgabe zu überzeugen, ein tendenziell repräsentativeres Gesamtbild der Verbrauchermeinung abgeben als andere.

E.2. Anbieter

Aufgrund ihres starken Einflusses auf die Verbraucherentscheidung sind Nutzerbewertungen für Anbieter von großer Bedeutung. Das gilt sowohl für Online-Händler wie z. B. einen Anbieter auf Amazon als auch für offline tätige Anbieter wie Ärzte oder Hotels, wenn deren Tätigkeit auf einer digitalen Plattform bewertet wird. Nutzerbewertungen sind für den einzelnen Anbieter bzw. seine Items auf der Plattform ein entscheidender Wettbewerbsfaktor im direkten Vergleich mit seinen Konkurrenten bzw. deren Items, während das Portal selbst vor allem ein allgemeines Interesse an einer möglichst hohen Zahl an Nutzerbewertungen hat.

⁷² Z. B. Frankfurter Allgemeine Sonntagszeitung, *Schummelei mit Sternen*, 06.10.2019.

E.2.1. Einfluss von Bewertungen auf den Umsatz der Anbieter

Im Rahmen der Sektoruntersuchung wurden alle ausgewählten Portale zu den Auswirkungen von Nutzerbewertungen auf ihren Geschäftserfolg befragt (s. Kapitel E.4.1.). Eine unmittelbare Befragung der auf einer Plattform vertretenen Anbieter von Produkten oder Dienstleistungen hat hingegen nicht stattgefunden. Allerdings existiert zu diesem Thema bereits eine umfangreiche wissenschaftliche Literatur, die den Zusammenhang zwischen Bewertungen und Geschäftserfolg hinreichend belegen und sogar noch spezifizieren kann. Ergänzend dazu wurde in der Befragung der Portale durch das Bundeskartellamt ermittelt, welche Bedeutung der Einbezug von Nutzerbewertungen beim Ranking der Suchergebnisse hat.

Nutzerbewertungen auf Online-Portalen gibt es seit Mitte der 1990er-Jahre, danach hat das Phänomen schnell an Bedeutung gewonnen. Wissenschaftliche Literatur hat sich mit der Quantifizierung und Spezifizierung der Effekte von Nutzerbewertungen insbesondere in den 2000ern intensiv befasst. An erster Stelle stand dabei die Frage, von besonderer Relevanz für die Marketingliteratur, in welcher Form und in welcher Stärke Nutzerbewertungen Auswirkungen auf die Umsätze des bewerteten Items haben.

Die 2016 veröffentlichte Metastudie von Babić Rosario et al.⁷³, die wesentliche Literatur zu dem Thema zusammenfassend auswertet, kann dabei einige übergreifende stilisierte Fakten zu Auswirkungen von Nutzerbewertungen aufzeigen. Die Metastudie beruht auf 96 einzelnen Studien aus über zehn Jahren und umfasst 40 Plattformen und 26 Produktkategorien in 11 Ländern. Die Autoren stellen fest, dass über alle Studien hinweg Bewertungen die Umsätze des bewerteten Items spürbar erhöhen. Die absolute Anzahl an Bewertungen beeinflusst den Umsatz dabei stärker als der Anteil positiver Bewertungen: Viele Bewertungen zu haben ist demnach wichtiger als gute Bewertungen zu haben. Der starke Einfluss der Anzahl an Bewertungen auf den Umsatz wird auf den „Mitläufereffekt“ zurückgeführt: Nutzer entscheiden sich eher für Items, für die sich bereits viele andere Nutzer entschieden haben. Das Risiko eines Fehlkaufs bzw. einer Fehlentscheidung soll dadurch reduziert werden. Dieser „Herdentrieb“ hat einen selbstverstärkenden Effekt, denn Items mit mehr Bewertungen werden eher ausgewählt und erhalten deshalb auch wieder mehr Bewertungen. Negative Bewertungen führen lt. Babić Rosario et al. nur in ausgewählten Fällen zu einer Verringerung der Umsätze. Meistens dominiert der Mengeneffekt, so dass mehr negative Bewertungen nicht zu weniger, sondern ggfs. sogar zu etwas höheren Umsätzen führen. Insgesamt ist der Einfluss von Bewertungen auf den Umsatz des bewerteten Produkts auf Handelsplattformen stärker als auf Bewertungsplattformen oder in sozialen Netzwerken, allerdings haben Bewertungen von Dienstleistungen gerade in sozialen Netzwerken einen sehr starken Einfluss. Stärker ist der Einfluss von Bewertungen auf Umsätze auch dann, wenn es sich um neue Produkte handelt.

Einzelstudien können zu diesen allgemeinen Ergebnissen noch einige Details beitragen und zeigen weitergehende Effekte für die betroffenen Märkte auf: Luca⁷⁴ untersucht den Einfluss positiver Bewertungen auf Yelp

⁷³ Rosario, Sotgiu, de Valck, Bijmolt, The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors, *Journal of Marketing Research*, Band 53, Ausgabe 3, S. 297-318, 2016.

⁷⁴ Luca, Reviews, Reputation, and Revenue: The Case of Yelp.com, Harvard Business School Working Paper 12-016.

auf den Umsatz von Restaurants am Beispiel von Seattle in den Jahren 2003 bis 2009. Zu Beginn dieses Zeitraums gab es noch keine relevante Bewertungsplattform, am Ende war Yelp die in Seattle mit Abstand am meisten genutzte Plattform, auf der zum damaligen Zeitpunkt 70 % aller Restaurants vertreten waren. Luca zeigt, dass eine um einen Stern bessere Bewertung bei *unabhängigen* Restaurants zu einer Umsatzsteigerung von etwa 9 % führen kann. Auf den Umsatz von Restaurantketten haben die Bewertungen auf Yelp hingegen keinen Einfluss. Dass Nutzer die Bewertungen auf Yelp als Informationsquelle heranziehen, um die Qualität unabhängiger Restaurants einschätzen zu können, und dass das Risiko, ein neues Restaurant zu besuchen, durch die Bewertungsplattform geringer wird, zeigt der Anstieg des Marktanteils unabhängiger Restaurants gegenüber Restaurantketten in Seattle durch die Verbreitung von Yelp als Bewertungsportal. Einen ähnlichen Effekt weisen Zhu et al.⁷⁵ nach, die die Auswirkungen von Nutzerbewertungen auf einer Bewertungsplattform für Videospiele auf die Umsätze mit diesen Spielen untersuchen: Sie zeigen, dass die Umsätze mit Spielen, die häufiger und besser bewertet werden, steigen, und dass dieser Effekt für weniger bekannte Spiele stärker ist. Auch hier wird deutlich, dass Nutzerbewertungen besonders wichtig für Produkte sind, deren Qualität der Verbraucher schwer einschätzen kann bzw. deren Kauf für ihn gefühlt mit einem höheren Risiko eines Fehlkaufs verbunden ist.

Der beschriebene selbstverstärkende Effekt wirkt allerdings nicht endlos, weil der Wert zusätzlicher Bewertungen mit steigender Zahl der Bewertungen abnimmt: Die absoluten Zahlen, die dazu in verschiedenen Studien und auch von Marktteilnehmern genannt werden, unterscheiden sich zwar; insgesamt herrscht aber Einigkeit darüber, dass die ersten fünf bis fünfzig Bewertungen einen sehr starken Effekt haben, dieser aber ab hundert und mehr Bewertungen deutlich abnimmt.

Die im Rahmen der Ermittlungen des Bundeskartellamts durchgeführte Portalbefragung hat insbesondere einen Aspekt bezüglich des Einflusses von Bewertungen untersucht, der in der wissenschaftlichen Literatur bisher weniger betrachtet, von Marktteilnehmern aber immer wieder als relevant hervorgehoben wurde. Es handelt sich um den indirekten Einfluss der Bewertungen auf Umsätze über den Umweg der Sichtbarkeit und Auffindbarkeit auf einem Portal, wenn der Portalbetreiber die Zahl bzw. Qualität der Bewertungen in seinem Ranking der Angebote berücksichtigt: Bei jeder Suche auf einem Portal werden dem Nutzer die Suchergebnisse in einer bestimmten Reihenfolge präsentiert. Ergebnisse, die weiter oben oder anderweitig prominenter platziert werden, werden vom Nutzer mit größerer Wahrscheinlichkeit ausgewählt. Ergebnisse, die erst auf Folgeseiten dargestellt werden, haben dagegen kaum eine Chance, wahrgenommen und ausgewählt zu werden. Um das Auffinden von Angeboten für den Nutzer möglichst komfortabel und erfolgreich zu gestalten und damit die Konversionsrate zu steigern, bemühen sich Portale, dem Nutzer auf den hervorgehobenen Rängen möglichst diejenigen Items anzuzeigen, die ihn am ehesten zufrieden stellen. Ein Kriterium, das dabei herangezogen werden kann, ist das Vorhandensein vieler bzw. positiver Nutzerbewertungen.⁷⁶ Items mit vielen Nutzerbewertungen werden häufiger ausgewählt und erhalten deshalb noch weitere Bewertungen. Dieser selbstverstärkende Effekt wird durch die Berücksichtigung von Bewertungen bei der Ergebnisdarstellung

⁷⁵ Zhu, Zhang, Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics, Journal of Marketing, Band 74, S. 133-148, 2010.

⁷⁶ Vgl. dazu auch Schmieder, *Zwielichtige Sternchen*, Süddeutsche Zeitung vom 04.05.2020, S. 18.

durch einen weiteren selbstverstärkenden Effekt ergänzt, da die bevorzugte Platzierung zusätzlich dazu führt, dass diese Items mit größerer Wahrscheinlichkeit ausgewählt werden. Das führt wiederum ebenfalls zu weiteren Bewertungen.

Der beschriebene Effekt kommt zwar regelmäßig, aber nicht auf allen Portalen zum Tragen, weil die Darstellung der Items nicht immer von Bewertungen abhängt. Bei den Portalen, die *Unternehmensbewertungen* darstellen, haben in der Befragung des Bundeskartellamtes etwa zwei Drittel angegeben, dass die Anzahl bzw. die Güte der Bewertungen die Reihenfolge der Suchergebnisse beeinflusse. Bei Portalen mit *Produktbewertungen* war es knapp die Hälfte. In beiden Kategorien ist bei vielen Portalen der Einfluss der Bewertungen nach eigenen Angaben jedoch nur sehr gering, so z. B. bei Amazon. Auch die Appstore-Betreiber gaben an, dass Sternebewertungen die Anzeige im jeweiligen Appstore beeinflussten, allerdings ebenfalls nur sehr geringfügig.

Insgesamt bestätigen die Ermittlungen des Bundeskartellamtes jedoch die in der wissenschaftlichen Literatur mittlerweile etablierte Erkenntnis, dass Items, die über mehr und bessere Bewertungen verfügen, eher vom Nutzer ausgewählt werden und dementsprechend mehr Umsätze haben.

E.2.2. Generierung von Nutzerbewertungen durch Anbieter

Da Bewertungen einen derart starken positiven Einfluss auf Umsätze aufweisen, ist es naheliegend, dass Anbieter versuchen, dieses Potential zu nutzen und möglichst viele (positive) Bewertungen für die eigenen Items zu erhalten. Wie oben dargestellt sind Bewertungen besonders wichtig für Items, die eher unbekannt sind und noch keine oder nur sehr wenige Bewertungen haben. Die Möglichkeiten, die den Anbietern zur Generierung von zusätzlichen Bewertungen zur Verfügung stehen, sind allerdings begrenzt. In einigen Fällen haben sie die Option, in Kooperation mit einem Portal den Nutzern gezielt Anreize zur Bewertungsabgabe zu setzen bzw. an einem vom Portal angebotenen Produkttesterprogramm teilzunehmen. Darüber hinaus können Anbieter Personen aus dem eigenen Umfeld um „Gefälligkeitsbewertungen“ bitten oder über Bewertungsmittler versuchen, Produkttester zu akquirieren, die ein Produkt kostenlos testen und im Gegenzug dafür eine Bewertung schreiben. Schließlich besteht für Anbieter grundsätzlich auch die Option, im großen Stil mehr oder weniger automatisiert positive Bewertungen erstellen zu lassen, also zu „kaufen“.

Bezüglich der Kooperation mit Portalen haben einige Handelsportale im Rahmen der Befragung angegeben, dass sie gelegentlich zusammen mit Anbietern Gewinnspiele veranstalten, bei denen für die Nutzer Anreize zur Bewertung von Artikeln einer bestimmten Marke gesetzt werden (z. B. „Nivea-Produkte bewerten und gewinnen“ (s. Kapitel D.1.3.)). Um Händlern gezielt die Möglichkeit zu geben, Bewertungen für Produkte zu erhalten, die noch keine oder erst wenige Bewertungen haben, bietet Amazon diesen in einigen Ländern

(nicht in Deutschland) das sog. Early Reviewer Program für bestimmte Produkte mit maximal fünf Bewertungen an.⁷⁷ Die Teilnahme an dem Programm ist für den Händler kostenpflichtig; der Nutzer, der eine Bewertung verfasst, erhält einen kleinen Geldbetrag bzw. Gutschein in Höhe von \$1 bis \$3. Während bei den Gewinnspielen und auch beim Early Reviewer Program dem Nutzer lediglich Anreize gesetzt werden, eine Bewertung abzugeben, suchen die Portale bei ihren eigenen Produkttesterprogrammen gezielt Nutzer, die ein konkretes, kostenlos zur Verfügung gestelltes Produkt testen und dann bewerten. Ein derartiges reguläres Produkttesterprogramm bietet unter den Portalen nach den Ermittlungen bisher nur Amazon mit Amazon Vine an.⁷⁸

Bei Vine erhalten ausgewählte Anbieter die Möglichkeit, Produkttests durch von Amazon ausgewählte Kunden durchführen zu lassen, die dabei als Produkttester fungieren. Vine-Bewertungen werden auf der Website von Amazon an gleicher Stelle und in gleicher Form wie die übrigen Nutzerbewertungen präsentiert, sind aber jeweils individuell eindeutig als Vine-Bewertung gekennzeichnet, so dass deren besonderes Zustandekommen für den Verbraucher erkennbar ist. Amazon hat das Programm im Dezember 2019 für Händler geöffnet, die die Teilnahmebedingungen erfüllen. Abgesehen davon, dass die Teilnahme an dem Programm grundsätzlich gewissen Beschränkungen unterliegt, kritisieren einige Marktteilnehmer, dass über das Vine-Programm generierte Nutzerbewertungen für die Hersteller bzw. Händler mit hohen Kosten verbunden seien.

Die Möglichkeiten der Anbieter, in Kooperation mit den Portalen Bewertungen zu generieren, sind begrenzt. Viele Anbieter versuchen deshalb, auf anderen Wegen Bewertungen für ihre Items zu generieren. Zunächst besteht dabei für Anbieter die Möglichkeit, gezielt Personen im eigenen Umfeld aufzufordern, das eigene Item zu bewerten. Insbesondere im Bereich der Produktbewertungen können Anbieter darüber hinaus über externe Bewertungsvermittler oder Gruppen in sozialen Netzwerken versuchen, Produkttester zu akquirieren, denen das betreffende Produkt dann kostenlos oder vergünstigt zur Verfügung gestellt wird und die im Gegenzug eine Bewertung verfassen sollen.

Bei den Gefälligkeitsbewertungen von Personen im eigenen Umfeld des Anbieters besteht aus Sicht des lesenden Verbrauchers keine Garantie, dass der Verfasser der Bewertung das zu bewertende Item überhaupt selbst genutzt, also z. B. ein Produkt ausprobiert oder ein Hotel besucht hat. Außerdem ist zu erwarten, dass mit der Bitte um die Bewertung mehr oder weniger explizit auch auf den Inhalt der Bewertung und die Zahl

⁷⁷ <https://www.amazon.com/gp/help/customer/display.html?nodetid=202094910> (abgerufen am 01.10.2020).

⁷⁸ Vine war ursprünglich nur für Produkte gedacht, die Amazon selbst als Verkäufer auf dem eigenen Marktplatz verkauft. In Amazons und in ihrem eigenen Interesse konnten die Hersteller dieser Produkte – die „Vendoren“ von Amazon – über Vine-Bewertungen den Absatz ihrer Produkte durch Amazon fördern. Da jedoch die anderen „Marktplatz-Händler“ und deren Produkte von Vine ausgeschlossen waren, sahen sich diese Händler und die jeweiligen Produkthersteller im Wettbewerb behindert. Im Verfahren B2-88/18, das hauptsächlich zur Verbesserung der allgemeinen Händler-Bedingungen von Amazon führte, hatte das Bundeskartellamt daher u. a. eine Öffnung des Vine-Programms für Marktplatzhändler, die Markenrechtsinhaber bzw. Bevollmächtigte sind, erwirkt. Vgl. *Bundeskartellamt*, Pressemitteilung v. 17.07.2019, https://www.bundeskartellamt.de/SharedDocs/Meldung/DE/Pressemitteilungen/2019/17_07_2019_Amazon.html (abgerufen am 01.10.2020).

der vergebenen Sterne Einfluss genommen wird. Dagegen ist bei Produkttestern grundsätzlich davon auszugehen, dass diese das Produkt tatsächlich erhalten haben, also über eigene Erfahrungen verfügen. Teilweise werden die Bewertungsverfasser dabei explizit um eine ehrliche Bewertung gebeten, teilweise wird aber auch hier verlangt, dass die Bewertung vier oder fünf Sterne enthält (vgl. Kapitel E.3.).

Eine explizite Kennzeichnung der so entstandenen Bewertungen durch den Anbieter ist in aller Regel nicht möglich, da die Portale häufig keine Darstellung incentivierter Bewertungen außerhalb ihrer eigenen Programme vorsehen. Im Rezensionstext selbst erfolgt häufig ebenfalls kein Hinweis auf die Incentivierung. Der Grund dafür kann einerseits darin liegen, dass nicht erkennbar sein soll, dass es sich um einen Produkttest handelt. Andererseits kann aber seitens des Anbieters auch die Befürchtung bestehen, dass Bewertungen, die einen entsprechenden Hinweis enthalten, vom Portal gelöscht werden.

Eine große Zahl von (positiven) Bewertungen kann auch in mehr oder weniger automatisierter Form erstellt werden, z. B. über Bots oder das „Kapern“ existierender Nutzer-Accounts. In diesen Fällen kommt die Bewertung unabhängig von einer tatsächlichen Nutzung des betreffenden Items zustande. Bei derart automatisiert erstellten Bewertungen ist außerdem davon auszugehen, dass sie – wenn der Bewertete sie selbst in Auftrag gegeben hat – grundsätzlich positiv sind. Weder der Anbieter des betreffenden Items noch der Ersteller dieser Art von Bewertungen dürften in dieser Situation ein Interesse daran haben, den Verbraucher über das Zustandekommen dieser Bewertungen aufzuklären und beide werden dementsprechend keine Kennzeichnung vornehmen.

Im Folgenden werden Bewertungen, die ohne eine eigene Erfahrung des Verfassers zustande gekommen sind, als **nicht-authentische Bewertungen** bezeichnet. Nicht-authentische-Bewertungen dürften v. a. bei Gefälligkeitsbewertungen oder bei automatisiert erstellten Bewertungen vorkommen und in aller Regel darüber hinaus inhaltlich manipuliert sein, da der Verfasser explizit um eine positive oder (im Falle der Bewertung eines Wettbewerbers) negative Bewertung gebeten wird. Bewertungen, bei denen der Verfasser zwar über eigene Erfahrung mit dem Item verfügt, das Produkt aber für einen Produkttest zur Verfügung gestellt bekommen hat, werden als **Produkttest-Bewertungen** bezeichnet. Wenn ein hoher finanzieller Anreiz in anderer Form erfolgt, handelt es sich um **incentivierte Bewertungen**. Wenn die Verfasser explizit zur Abgabe einer positiven (oder negativen) Bewertung aufgefordert werden, handelt es sich um **manipulierte Bewertungen** (s. Abbildung 15 in Kapitel F.).

Bei Unternehmen oder Dienstleistungen sind die Möglichkeiten der Anbieter, Bewertungen über Tester zu generieren, begrenzt. Ein Test-Besuch beim Arzt oder ein kostenloser Test-Urlaub im Hotel dürfte in der Praxis schwieriger umzusetzen sein, als der Versand eines Produkts. Dafür kann es bei Unternehmen oder Dienstleistungen einfacher sein, um Gefälligkeitsbewertungen zu bitten, weil nicht ein konkretes Produkt eines Händlers bewertet werden muss, sondern nur ein einzelner Anbieter wie z. B. ein Restaurant oder Hotel. Auch im Bereich der Unternehmensbewertungen ist folglich das Problem nicht-authentischer Bewertungen, die ohne tatsächliche Nutzung der betreffenden Leistung abgegeben wurden, bzw. manipulierter Bewertungen, deren Inhalt beeinflusst wurde, gleichermaßen relevant.

Zwei wissenschaftliche Studien zu dem Thema versuchen mit Hilfe von Indizien, Rückschlüsse auf das Vorhandensein nicht-authentischer bzw. manipulierter Bewertungen zu ziehen. Sie beziehen sich dabei auf Hotels bzw. Restaurants: Mayzlin et al.⁷⁹ verglichen Bewertungen auf einer Plattform mit einem transaktionsbasierten, geschlossenen Bewertungssystem und einer Plattform mit einem offenen Bewertungssystem aus dem Jahr 2011 für 50 US-amerikanische Städte. Sie stellten fest, dass kleine, inhabergeführte Hotels auf der Plattform mit dem offenen Bewertungssystem relativ mehr und mehr positive Bewertungen aufwiesen als auf der Plattform mit dem geschlossenen System. Hotels, die enge Wettbewerber von kleinen, inhabergeführten Hotels sind, hatten auf der offenen Plattform ebenfalls relativ mehr, allerdings mehr negative Bewertungen als auf der geschlossenen Plattform. Die Autoren schlossen daraus, dass es sich zumindest bei diesen zusätzlichen um nicht-authentische und manipulierte Bewertungen handelte. Kleine, inhabergeführte Hotels profitieren mehr als Hotelketten von zusätzlichen Bewertungen und haben deshalb auch einen höheren Anreiz, für entsprechende Bewertungen „zu sorgen“. Luca et al.⁸⁰ beschrieben die gleichen Effekte bei Restaurants auf dem Bewertungsportal Yelp, ebenfalls in den USA und für das Jahr 2014: Kleine inhabergeführte Restaurants wiesen dort relativ mehr positive Bewertungen auf als Ketten. Von kleinen, inhabergeführten Restaurants benachbarte Restaurants wiesen dagegen relativ mehr negative Bewertungen auf. Positive Bewertungen, bei denen der Verdacht bestand, dass sie nicht-authentisch und manipuliert sind, erschienen zudem eher dann, wenn ein Restaurant direkt davor negative Bewertungen erhalten hatte. Zur Identifikation solcher Bewertungen griffen die Autoren auf die Ergebnisse des Algorithmus von Yelp zurück sowie auf Ermittlungen, die Yelp durchgeführt hatte, um gegen „fake reviews“ vorzugehen. Dabei hatte Yelp auf Aufforderungen von Restaurants, Bewertungen abzugeben, geantwortet, um herauszufinden, welche Restaurants solche Bewertungen in Auftrag geben, und die entsprechenden Restaurants dann auf ihrer Website mit einem entsprechenden Hinweis versehen.

Einige der Portale, die das Bundeskartellamt befragt hat, haben dargelegt, dass sie große Anstrengungen unternehmen, gegen nicht-authentische bzw. manipulierte Bewertungen vorzugehen und dabei auch Möglichkeiten nutzen, die zum Zeitpunkt der zitierten Studien in dieser Form noch nicht zur Verfügung standen. Es gibt aber auch Portale, deren Filtermaßnahmen nach wie vor nicht ausreichend sind. Deshalb ist auch heute und in Deutschland davon auszugehen, dass viele Nutzerbewertungen, die Verbraucher lesen und an denen sie sich orientieren, auf Initiative des jeweiligen Anbieters als Gefälligkeit durch Freunde oder Bekannte, über Produkttester oder in automatisierter Form zustande gekommen sind (s. Kapitel B.). Eine entsprechende Kennzeichnung der jeweiligen Nutzerbewertungen erfolgt in der Regel – auch mangels entsprechender Möglichkeiten auf den Portalen – überhaupt nicht bzw. höchstens im Rezensionstext, so dass der Nutzer nicht oder nur mühsam erkennen kann, dass es sich ggfs. um eine manipulierte oder nicht-authentische Bewertung handelt.

⁷⁹ Mayzlin, Dover, Chevalier, Promotional Reviews: An Empirical Investigation of Online Review Manipulation, *American Economic Review*, Band 104, Ausgabe 8, S. 2421-2455, 2014.

⁸⁰ Luca, Zervas, Fake it till you make it: Reputation, Competition and Yelp Review Fraud, *Management Science*, Band 62, Ausgabe 12, S. 3393-3672, 2016.

E.2.3. Wesentliche Erkenntnisse

Eine höhere Anzahl von Bewertungen für einen Item wirkt sich positiv auf den Umsatz des Anbieters mit diesem Item aus. Dabei ist es nicht unbedingt notwendig, besonders gute Bewertungen zu haben. Durch den selbstverstärkenden Effekt werden Items, die bereits Bewertungen haben, vermehrt gekauft und erhalten weitere Bewertungen. Items ohne Bewertungen gehen tendenziell weiterhin leer aus. Der Effekt verstärkt sich durch die Vorgehensweise mancher Portale, Items mit vielen Bewertungen im Ranking besser zu platzieren als solche mit wenigen. Der Einfluss der Zahl der Bewertungen auf die Verbraucherentscheidung und damit den Umsatz nimmt allerdings ab einer gewissen Größenordnung wieder ab. Einen besonders starken Einfluss hat die Zahl der Bewertungen bei unbekanntem Produkten oder unbekanntem Dienstleistern, die der Verbraucher ansonsten nur schwer einschätzen kann.

Aufgrund der fundamentalen Bedeutung von Nutzerbewertungen insbesondere im Online-Handel und der begrenzten Möglichkeiten der Anbieter, über die Plattformen offiziell Nutzerbewertungen zu generieren, bitten Anbieter Bekannte um Gefälligkeitsbewertungen, setzen Produkttester ein oder lassen in großem Stil positive Bewertungen produzieren. Grundsätzlich zu unterscheiden ist dabei, ob der Bewertung tatsächlich eine Nutzung zugrunde lag oder nicht (Authentizität), ob auf den Inhalt der Bewertung Einfluss genommen wurde oder nicht (Manipulation) und ob die Abgabe der Bewertung mit einem Anreiz verbunden war oder nicht (Incentivierung/Produkttest). Wenn derart zustande gekommene Bewertungen nicht gekennzeichnet und damit identifiziert werden (können), hat der Verbraucher keine Möglichkeit zu erkennen, warum eine Bewertung abgegeben wurde bzw. welche Motive der Bewertung zugrunde lagen. Anbieter, die versuchen, echte und unbeeinflusste Bewertungen über Produkttester zu erhalten, können sich zudem nicht von solchen Anbietern abgrenzen, die garantierte 5-Sterne-Bewertungen kaufen. Es besteht die Gefahr einer Irreführung der Verbraucher und eines grundsätzlichen Vertrauensverlustes in Bewertungen (s. Kapitel F.1. bis F.4.).

E.3. Bewertungsvermittler

Bewertungsvermittler sind Dienstleister, die die Vermittlung von Nutzerbewertungen für Dritte anbieten. Gemeinsam ist den typischen Bewertungsvermittlern im hier verstandenen Sinne, dass sie (echten) Privatpersonen Incentives dafür anbieten, dass diese bestimmte Leistungen, die sie selbst in Anspruch genommen haben, bewerten. Die Incentivierung besteht i. d. R. in der Möglichkeit, die zu bewertende Leistung kostenlos oder rabattiert zu erhalten, gelegentlich auch in einer finanziellen Vergütung in Form von Geldbeträgen oder Gutscheinen. Die Rolle der Bewertungsvermittler ist in der folgenden Graphik dargestellt: Beauftragt und bezahlt werden die Vermittler von den Anbietern. Ihre Funktion ist es, aus der Gruppe der Verbraucher Produkttester zu akquirieren, die dann die dafür vorgesehenen Leistungen testen und bewerten. Diese privaten Produkttester sollen dann ihre Bewertung i. d. R. auf dem Portal einstellen, auf dem der Anbieter die zu testende Leistung anbietet, und übersenden diese zusätzlich auch an den Bewertungsvermittler.

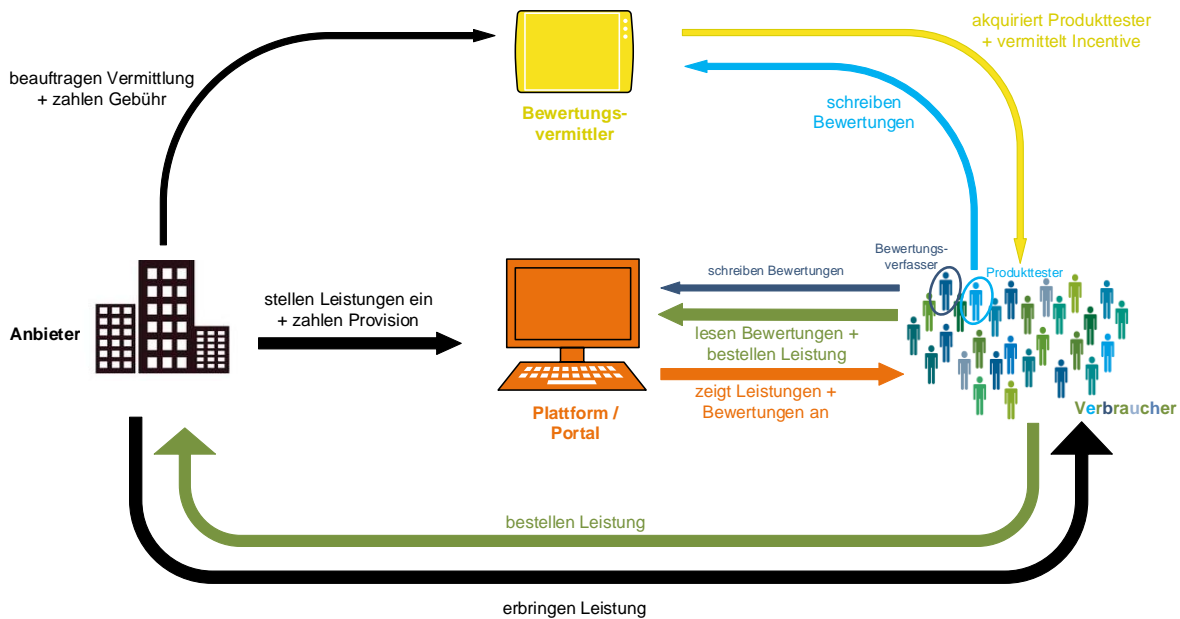


Abbildung 11: Bewertungsvermittler als zusätzliche Marktteilnehmer bei Nutzerbewertungen

Von dem hier betrachteten Ansatz der Bewertungsvermittlung grundsätzlich zu unterscheiden ist die mehr oder weniger automatisierte Erstellung einer Vielzahl von Bewertungen, ohne dass der Verfasser das Produkt oder die Leistung tatsächlich genutzt hat (z. B. durch sog. „Bots“). Nicht im Fokus steht hier zudem der (paketweise) Verkauf von Followern oder Likes für (soziale) Netzwerke und Plattformen wie Facebook, Instagram oder YouTube, die ebenfalls häufig automatisiert erstellt werden. Besondere Systeme zur Generierung von Bewertungen sind auch die sog. „Communities“ oder „Nutzer-Foren“, in denen die Mitglieder insbesondere aufgrund ideeller Anreize („Gamification“) Bewertungen zu bestimmten Themen abgeben. Schließlich sind von den nachstehenden Ausführungen zu Bewertungsvermittlern auch *professionelle* Produkttester abzugrenzen, die für spezielle Organisationen oder Publikationen hauptberuflich eine bestimmte Art von Produkten oder Unternehmen testen (z. B. Restaurantkritiker, Stiftung Warentest).

Die Ermittlungen zu Bewertungsvermittlern haben gezeigt, dass deren Geschäft im Wesentlichen auf zwei Wegen stattfindet: Der eine Weg führt über Unternehmen, die als Dienstleister die Vermittlung von Bewertungen anbieten sowie teilweise darüber hinaus auch noch andere Bereiche des Online-Marketings. Diese Unternehmen sind in der Regel als GmbH, UG, Ltd. oder LLC organisiert und haben häufig ihren Sitz im Ausland. Der andere Weg verläuft über Gruppen in (sozialen) Netzwerken wie Facebook, WhatsApp und Telegram, wo Anbieter bzw. Bewertungsvermittler mit potenziellen Produkttestern in Kontakt treten. Die formelle Organisation dieser Gruppen erfolgt allerdings häufig über Privatpersonen. Teilweise vermischen sich die beiden Entstehungswege, wenn beispielsweise Produkttester gezielt über (soziale) Medien gesucht werden, die restliche Abwicklung aber außerhalb des Netzwerks erfolgt.

Auf den Seiten der jeweiligen Portale sind die über Bewertungsvermittler incentiviert zustande gekommenen Bewertungen für den Verbraucher nicht oder nur schwer als solche zu erkennen. Eine offizielle Kennzeichnung der Bewertung als Produkttest durch Dritte (also entweder den Verfasser oder den Vermittler) ist auf den befragten Portalen nicht möglich. Es bleibt daher allein die Möglichkeit, im Text der Rezension darauf

hinzuweisen, dass es sich um eine incentivierte Bewertung handelt, z. B. durch einen Hinweis wie „Ich habe das Produkt kostenlos erhalten, um es zu testen“ am Beginn des Rezensionstextes.

Auch dieser Hinweis erfolgt jedoch häufig nicht. Ein Grund dafür mag sein, dass für den Leser nicht erkennbar sein soll, dass die Bewertung incentiviert erfolgt ist, weil der Leser Bewertungen, die intrinsisch motiviert zustande gekommen sind, mutmaßlich mehr vertraut. Ein weiterer Grund kann die Befürchtung sein, dass Bewertungen, die einen entsprechenden Hinweis im Rezensionstext enthalten, vom Portal gelöscht werden. Denn eine Reihe von Portalen schließt die Abgabe incentivierter Bewertungen in ihren AGBs grundsätzlich aus. Amazon zum Beispiel weist darauf in seinen Community-Richtlinien hin und gibt an, entsprechende Bewertung zu löschen.⁸¹ Diese Vorgaben haben in den letzten Jahren allerdings nicht verhindert, dass durch die Tätigkeit gewerblich tätiger Bewertungsvermittler incentivierte Bewertungen auf Portalen abgegeben und dargestellt werden.

E.3.1. Vermittlung von Bewertungen über Unternehmen

Um die Geschäftsmodelle und die Interessenlagen der unterschiedlichen Bewertungsvermittler besser einschätzen zu können, hat das Bundeskartellamt eine Reihe von Unternehmen, die entsprechende Dienstleistungen anbieten, um Stellungnahme zu verschiedenen Themen gebeten. Aufgrund der z. T. deutlichen, auch öffentlichen Kritik an den Bewertungsvermittlern und der rechtlichen Grauzone, in der sie sich teilweise bewegen (siehe nachfolgende Abschnitte), ist eine Kontaktaufnahme zu den Bewertungsvermittlern schwierig. Viele ändern regelmäßig ihren Namen bzw. treten unter unterschiedlichen Bezeichnungen auf, und viele haben ihren Unternehmenssitz ins (außereuropäische) Ausland verlegt. Alleine in fünf Fällen, davon drei in Deutschland, kamen die Anschreiben des Bundeskartellamtes als „unbekannt“ zurück, obwohl sie an die aus dem Impressum ersichtliche offizielle Postadresse des jeweiligen Unternehmens versandt worden waren. Auch die Händler bzw. Hersteller, die die Dienste der Bewertungsvermittler in Anspruch nehmen, wollen häufig nicht genannt werden. Dennoch haben einige Bewertungsvermittler eine schriftliche Stellungnahme abgegeben oder standen dem Bundeskartellamt sogar für ein persönliches Gespräch zur Verfügung.

Diese Ermittlungen haben gezeigt, dass sich die Geschäftsmodelle der Bewertungsvermittler in Bezug auf drei zentrale Punkte voneinander unterscheiden:

Unterschiede gibt es dabei zunächst in Bezug auf die Abläufe von der Beauftragung bis zu Einstellung einer Bewertung: In der Regel können die registrierten Bewertungsverfasser (bzw. Produkttester) auf der Website des Vermittlers sehen, welche Produkte aktuell zur Bewertung angeboten werden und sich dafür melden oder bewerben. Teilweise erhalten die Tester das Produkt dann (ggfs. über den Vermittler) direkt vom Händler/Hersteller, teilweise müssen sie das Produkt selbst erwerben und erhalten dafür entweder einen Gutschein oder bekommen (teilweise) den Kaufpreis erstattet. Die Erstattung erfolgt bei manchen Vermittlern

⁸¹ <https://www.amazon.de/gp/help/customer/display.html?nodeId=201929730>, https://www.amazon.de/gp/help/customer/display.html/ref=hp_bc_nav?ie=UTF8&nodeId=G3UA5WC5S5UUKB5G
(beide abgerufen am 01.10.2020).

im Vorhinein, häufig müssen die Tester dafür aber selbst in Vorkasse gehen. Der selbständige Kauf bei einer Plattform ist für die spätere Verwendung der Bewertung insbesondere dann wichtig, wenn – wie bei Amazon – Wert auf den Hinweis „verifizierter Kauf“ gelegt wird. Einige Vermittler geben den Bewertungsverfassern konkrete Vorgaben dazu, auf welchen Plattformen oder Portalen die Bewertungen zu veröffentlichen sind, andere lassen die Abgabe einer Bewertung vollständig offen oder fordern lediglich die Einstellung der Bewertung auf der eigenen Website.

Des Weiteren spielt die Frage der Kennzeichnung derart incentivierter Bewertungen für die Bewertungsvermittler eine Rolle. Hierzu wurde teilweise vorgetragen, dass die Verfasser dazu „angehalten“ seien, in ihrem Bewertungstext darauf hinzuweisen, dass sie dafür eine Vergütung oder ein kostenloses Produkt erhalten haben. Kritisiert wurde von verschiedenen Bewertungsvermittlern, dass einige Plattformen versuchen, Bewertungen mit solchen Hinweisen zu löschen, gleichzeitig aber keine Alternative für eine offizielle Kennzeichnung anbieten. Mehrere Bewertungsvermittler haben in diesem Zusammenhang betont, dass insbesondere Amazon die Abgabe incentivierter Bewertungen durch verschiedene Maßnahmen gezielt verhindere, um das eigene – für Händler teure – Bewertungsprogramm „Vine“ oder das „Early Reviewer Program“ zu fördern, dessen Mitglieder ebenfalls kostenlose Produkte zum Testen erhielten. Auch Holidaycheck belohne seine Nutzer teilweise selbst mit Boni, wenn diese eine Bewertung abgeben.

Ein wesentliches Unterscheidungskriterium ist schließlich die Frage, inwieweit der Bewertungsvermittler auf den Inhalt bzw. die Qualität der abgegebenen Bewertungen Einfluss nimmt. Während einige Unternehmen mehr oder weniger deutlich auf ihrer Homepage, in den Richtlinien oder in Schulungs-Videos darauf hinweisen, dass von den Verfassern „positive“ Bewertungen erwartet werden, betonen andere Vermittler, dass die Verfasser hinsichtlich der Zahl der Sterne und des Bewertungstextes völlig frei sind und lediglich gewisse Rahmenbedingungen bzgl. Produktbezug und Textlänge einhalten müssen. Die Beeinflussung kann auch darin bestehen, dass Produkttester, die keine 5-Sterne-Bewertung auf Amazon abgeben, keine weiteren Aufträge erhalten. Einer der befragten Bewertungsvermittler trug dazu vor, dass incentivierte Bewertungen zwar im Durchschnitt um 0,7 Sterne besser ausfallen, sich aber hinsichtlich Länge und Qualität des Bewertungstextes nicht nennenswert von nicht-incentivierten Bewertungen unterscheiden. Eine Incentivierung führe in erster Linie dazu, dass mehr Bewertungen abgegeben würden. Ein anderer Bewertungsvermittler hat berichtet, dass seine Produkttester sogar besonders kritisch seien. Mehrfach wurde von Seiten der Vermittler darauf hingewiesen, dass die Incentivierung von Bewertungen zu einem realistischeren Bewertungsdurchschnitt führe, da insbesondere Nutzer mit positiven oder mittelmäßigen Erfahrungen dadurch eher motiviert würden, überhaupt eine Bewertung zu schreiben. Dies sei auch deshalb erforderlich, weil das Gesamtbild der Bewertungen ansonsten durch die mangelhafte Kontrolle bzw. die vorgegebenen Algorithmen der Plattformen verzerrt würde. Einige Bewertungsvermittler haben auch darauf hingewiesen, dass sie selbst bereits Maßnahmen durchführen, um missbräuchliche Bewertungen zu verhindern.

Je nach Ausgestaltung des Geschäftsmodells in den genannten drei Punkten kann die Vermittlung incentivierter Bewertungen durch Unternehmen und deren Einsatz durch den betreffenden Händler mit verschiedenen rechtlichen Fragen verbunden sein, was zuletzt auch Gegenstand diverser gerichtlicher Auseinandersetzungen war. Betroffen sind davon die meisten großen Portale, mit Abstand am häufigsten angeboten werden Bewertungen für Amazon. Wenn bei derartigen Angeboten eine feste Zahl von Amazon-Bewertungen

gekauft werden kann (ggfs. verbunden mit der mehr oder weniger expliziten Zusage von fünf Sternen), handelt es sich mit hoher Wahrscheinlichkeit um unseriöse Anbieter, da für echte Produkttests eine solche Garantie unter Berücksichtigung der Amazon-Richtlinien nicht ausgesprochen werden könnte. Die Preise für eine Amazon-Bewertung liegen im Bereich von 15 bis 30 Euro pro Bewertung, je nach Vermittler und bestellter Menge. Nach Angaben von Marktteilnehmern lassen sich allein mit der Vermittlung von Bewertungen für den deutschen Amazon-Marktplatz sechs- bis siebenstellige Jahresumsätze erzielen.

Vier der befragten Portale, Amazon, Holidaycheck, Jameda und Mobile.de gehen durch Abmahnungen und Klagen vor Gericht teilweise bereits seit vielen Jahren aktiv gegen die Vermittlung von Bewertungen vor. Dabei haben sie bereits eine ganze Reihe von Unternehmen, die als Bewertungsvermittler tätig sind, sowie deren Geschäftsführer abgemahnt bzw. verklagt. Die Portale waren dabei überwiegend erfolgreich; einige Verfahren sind noch nicht abgeschlossen.

Amazon hat bisher die meisten Verfahren angestrengt und dabei bereits gegen eine ganze Reihe von Bewertungsvermittlern und deren Geschäftsführer einstweilige Verfügungen bei den Landgerichten in Frankfurt, Hamburg und München erwirkt, die den Betroffenen ihre Geschäftstätigkeit als Bewertungsvermittler für Bewertungen auf Amazon untersagen. In einigen dieser Fälle hatten sich die Betroffenen daraufhin mittels rechtsverbindlicher Abschlusserklärungen wie von Amazon gefordert dazu verpflichtet, ihre Vermittlungstätigkeit für (nicht-gekennzeichnete) Amazon-Nutzerbewertungen aufzugeben. In der Mehrzahl der Fälle haben die Beklagten gegen die einstweilige Verfügung Berufung eingelegt. In zwei Fällen wurde die Berufung zurückgewiesen, in drei weiteren Fällen wurde sie zurückgenommen. Holidaycheck hat ebenfalls mehrere Verfahren geführt. Beispielhaft sei hier genannt, dass das Landgericht München I einen Bewertungsvermittler und dessen Geschäftsführer auf eine entsprechende Klage von Holidaycheck hin dazu verurteilte, es zu unterlassen, im geschäftlichen Verkehr gegenüber Dritten den Erwerb von Bewertungen für Holidaycheck anzubieten und/oder zu bewerben.⁸² Neben Bewertungsvermittlern hat Holidaycheck auch Anbieter verklagt, die Bewertungsvermittler mit dem Erstellen von Bewertungen beauftragt hatten. Auch Jameda geht regelmäßig juristisch gegen Bewertungsvermittler vor.⁸³ Außerdem werden bei Auffälligkeiten die Bewertung eines Arztes mit einem Warnhinweis versehen (s. D.2.5.). Mobile.de schließlich hat erreicht, dass das Landgericht Hamburg auf entsprechende Klagen hin mehreren Bewertungsvermittlern und deren Geschäftsführern die Vermittlung positiver Bewertungen untersagt hat.⁸⁴

In Reaktion auf die Gerichtsverfahren sowie Richtlinien-Änderungen der Portale nehmen Bewertungsvermittler allerdings immer wieder Veränderungen ihrer Geschäftsmodelle vor. In den Stellungnahmen und Gesprächen im Rahmen der Sektoruntersuchung haben verschiedene Bewertungsvermittler betont, dass eine legale oder richtlinienkonforme Vermittlung von Produkttestern vor diesem Hintergrund kaum möglich sei: Solange

⁸² LG München I, Urt. v. 28.03.2019, Az. 3 HK O 14073/18 (nicht veröffentlicht).

⁸³ LG München I, Beschl. v. 31.08.2017, Az. 1 HK O 12781/17 (nicht veröffentlicht).

⁸⁴ LG Hamburg, Beschl. v. 14.05.2018, Az. 327 O 193/18 (nicht veröffentlicht) u. a.

es „schwarze Schafe“ gäbe, die den Anbietern zu niedrigen Preisen 5-Sterne-Bewertungen auf Amazon garantieren, wären seriöse Vermittler von unbeeinflussten Produkttestern ohne Bewertungsgarantie für die Anbieter nicht attraktiv.

E.3.2. Vermittlung von Bewertungen über Gruppen in (sozialen) Netzwerken

Parallel zur Vermittlung von Bewertungen über Unternehmen gibt es in den (sozialen) Netzwerken Gruppen, in denen die Vermittlung von Produkttests stattfindet. Nach den Ermittlungen erfolgt dies in signifikantem Umfang auf Facebook, Telegram und WhatsApp.

Eine von einer deutschen Facebook-Seite aus durchgeführte Suche nach Gruppen mit den Stichworten „Produkttest“ und „Amazon“ ergab eine Vielzahl von Treffern, darunter u. a. „Amazon DE 5 stars review“, „Amazon DE Produkttester Deutschland“, „Amazon FBA Verkäufer Deutschland“, „Amazon Free Product Tester Group“, „Amazon Germany reviewer“, „Amazon Produkttester“, „Amazon Produkttester Germany/DE“, „Amazon Reviews Count (Refund After Review)“, „AMZ-Hunter.de Die Deal Gruppe“ oder „Produkttester DE Amazon Deutschland“. Auch ohne die Eingabe von „Amazon“ in die Suchmaske ist der klare Schwerpunkt der Suchergebnisse auf Amazon bezogen. Beim Gruppen-Profilbild wird durch die Verwendung des Amazon-Schriftzugs oder –Logos, meist in leicht veränderter Form, ebenfalls häufig auf Amazon angespielt. Die nähere Betrachtung von 16 ausgewählten Gruppen, bei denen vom Namen her eindeutig deutsche Produkttester für auf Amazon Marketplace angebotene Produkte gesucht werden, hat gezeigt, dass diese im Durchschnitt jeweils 6300 Mitglieder und täglich ca. 60 neue Nachrichten hatten. In den meisten Fällen gibt es klare Hinweise, dass das Profil des Gruppenadministrators nicht das Hauptprofil einer realen Person ist. Im Wesentlichen das gleiche Bild ergibt sich bei Telegram, wo ebenfalls eine Gruppensuche mit den entsprechenden Stichworten zu einer längeren Liste von entsprechenden Gruppen führt.⁸⁵ Bei WhatsApp ist eine Gruppensuche nicht möglich, trotzdem existieren dort ebenfalls Bewertungsgruppen. Die folgende Abbildung zeigt beispielhaft zwei Screenshots von Produkttestergruppen auf Telegram und WhatsApp.

⁸⁵ Zur Vermittlung von Produkttests insbesondere von chinesischen Herstellern über Telegram siehe auch: Manager Magazin, *Amazon – Außer Kontrolle*, 20.03.2020, S. 56 ff. sowie Die Zeit, *Tausche Sterne gegen Ware*, 16.04.2020, S. 22. Zur vergleichbaren Situation in Großbritannien siehe: Financial Times, *Amazon deletes 20,000 reviews after evidence of profits for posts*, 04.09.2020.

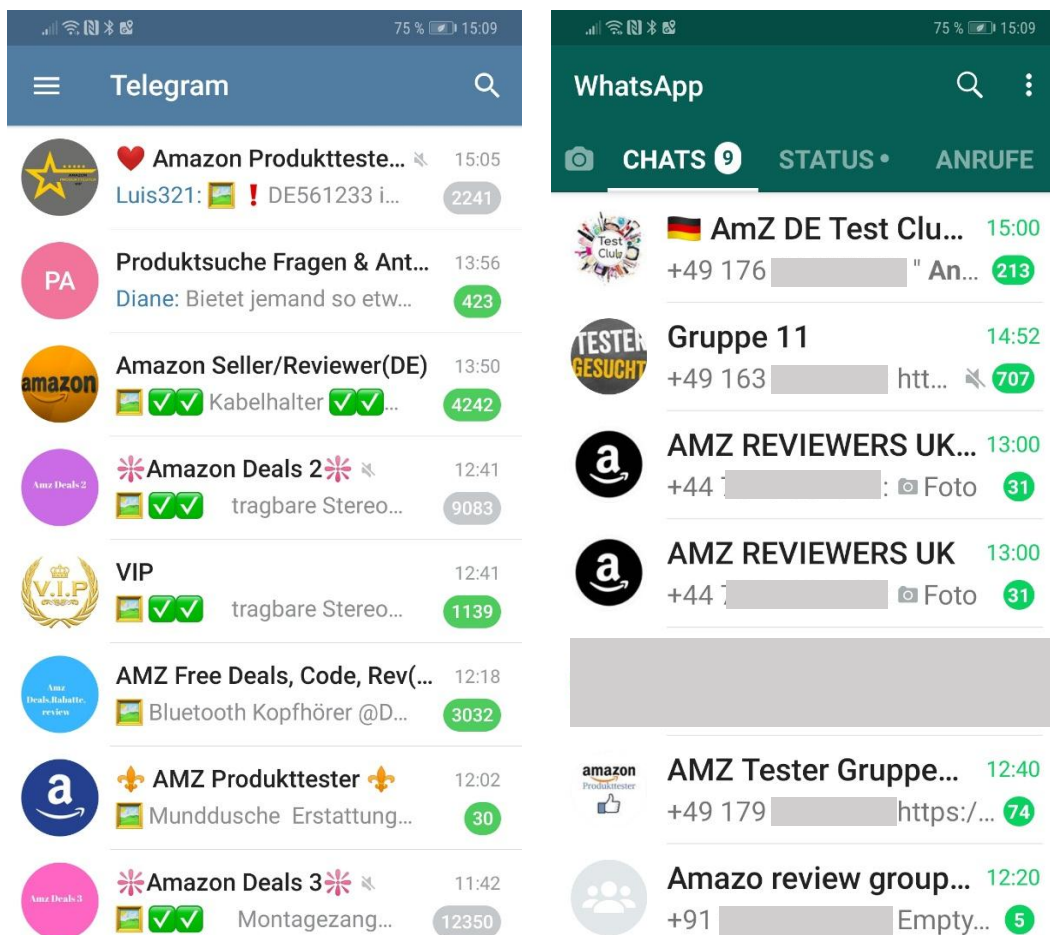


Abbildung 12: Screenshots von Produkttestergruppen auf Telegram und WhatsApp (Stand: Januar 2020)

Darüber hinausgehende Ermittlungen hat das Bundeskartellamt in diesem Bereich nicht durchgeführt. Bei den Gruppenadministratoren handelt es sich möglicherweise nicht um dem Bundeskartellamt gegenüber auskunftspflichtige Unternehmen, sondern um Privatpersonen. Es ist anzunehmen, dass Anbieter die Gruppenadministratoren ebenfalls bezahlen, analog zur Bezahlung, die fällig wird, wenn sie die Vermittlungsdienstleistung eines Unternehmens in Anspruch nehmen. Möglicherweise stehen auch als Bewertungsvermittler tätige Unternehmen, wie sie im vorigen Abschnitt beschrieben wurden, hinter den Administratoren bzw. deren Gruppen. Die Befragung bei den Portalbetreibern hat ergeben, dass sie bislang keine gerichtlichen Verfahren angestrengt haben, um gegen Bewertungsgruppen in Netzwerken vorzugehen. Insofern könnten diese Gruppen für Unternehmen, die als Bewertungsvermittler tätig sind, eine Umgehungsmöglichkeit darstellen, um weiterhin Produkttests zu vermitteln, ohne darauf auf ihrer Webseite im Internet aufmerksam machen zu müssen.

Bezüglich der Abläufe funktionieren die verschiedenen Bewertungs-Gruppen in (sozialen) Netzwerken jeweils in ähnlicher Weise: In der Regel muss der Produkttester zunächst das Produkt bei Amazon kaufen. Die Bewertung soll dann drei bis fünf Tage später auf Amazon abgegeben werden. Nachdem die Produkttester einen Screenshot ihrer Bewertung an den Vermittler oder Händler geschickt haben, bekommen sie das entsprechende Produkt erstattet und in Einzelfällen sogar noch ein paar Euro zusätzlich.

Eine Kennzeichnung der Bewertungen ist zumindest in den Gruppenchats kein Thema. Es ist daher davon auszugehen, dass in aller Regel keine Kennzeichnung der dadurch zustande gekommenen Bewertungen erfolgt. Aus den Chatverläufen ergibt sich, dass den Teilnehmern bewusst ist, dass Amazon versucht, die Bewertungen der Produkttester als solche zu identifizieren und zu löschen. Es werden daher auch Tipps gegeben, wie man vermeiden kann, dass Amazon Bewertungen löscht, z. B. indem man Produkte gleicher Marken nicht am gleichen Tag vom gleichen Account kauft.

Einfluss auf den Rezensionstext wird im Rahmen der Chatverläufe in der Regel nicht genommen, dafür wird teilweise vorgegeben, dass die Bewertung fünf Sterne betragen muss, z. B. mit Hinweisen wie in der folgenden Abbildung.

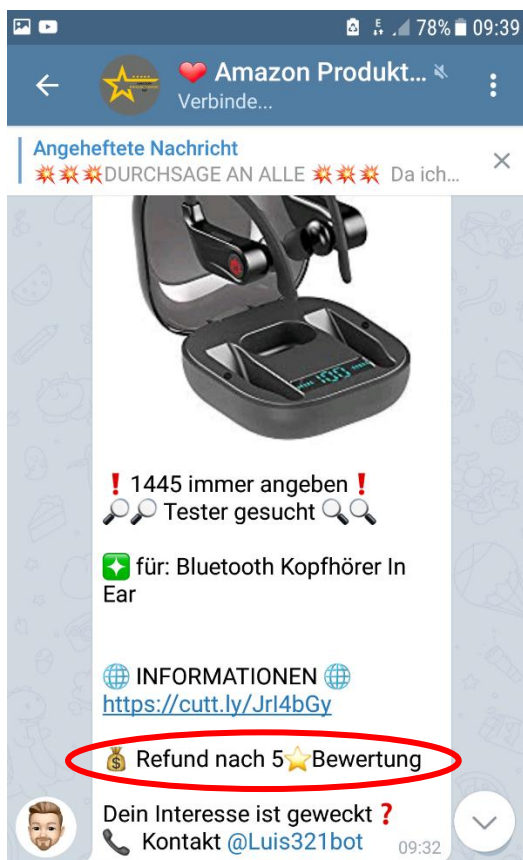


Abbildung 13: Aufforderung zur Vergabe von fünf Sternen auf Telegram (Stand: Februar 2020)

E.3.3. Wesentliche Erkenntnisse

Die Tätigkeit der Bewertungsvermittler basiert auf dem Wunsch der Anbieter nach guter Positionierung und der geringen Eigenmotivation der Verbraucher bei der Verfassung von Bewertungen. Gleichzeitig stehen den Bewertungsvermittlern jedoch das Eigeninteresse der Portale an den Bewertungen entgegen sowie die öffentlichen Vorwürfe wegen „gekaufter“ Bewertungen. Die Ermittlungen im Rahmen der Sektoruntersuchungen haben bestätigt, dass es zahlreiche Vermittlungen von Produkttests gibt. In der Regel werden die dadurch generierten Bewertungen auf den Portalen nicht als Produkttests gekennzeichnet; es handelt sich daher um verdeckte Produkttest-Bewertungen (s. Abb. 16). Die Ermittlungen haben aber auch gezeigt, dass das Bild

der Bewertungsvermittler deutlich differenzierter ist, als es die übliche Pauschal-Verurteilung zeichnet. Unterschiede in der Seriosität der Bewertungsvermittler zeigen sich bei Abläufen, Kennzeichnung und Einflussnahme: So gibt es durchaus Vermittler, die ihre Abläufe so strukturieren, dass kein vermeintlich „verifizierter Kauf“ stattfindet und der Produkttester für den Auftraggeber anonym bleibt, die auf die notwendige Kennzeichnung von Produkttests hinweisen und darauf achten, dass die Bewertung unbeeinflusst erfolgt. Diese seriösen Vermittler haben allerdings aus Sicht ihrer Kunden (also der Anbieter, Händler bzw. Hersteller) den Nachteil, keine fünf Sterne garantieren zu können. Außerdem werden alle Vermittler zunehmend in der Ausübung ihrer Geschäftstätigkeit durch die restriktive Praxis von Portalen wie Amazon gegenüber externen Produkttests eingeschränkt. Amazon hat zudem begonnen, mit Amazon Vine das wichtige Geschäft mit den Bewertungen immer weiter selbst zu übernehmen, was geeignet ist, auch die korrekt agierenden alternativen Bewertungsvermittler aus dem Markt zu verdrängen und insofern zumindest kartellrechtlich Fragen aufwirft. Angesichts der großen Bedeutung, die authentische und repräsentative Nutzerbewertungen für Verbraucher haben, sollte es für Händler und Vermittler auch in Zukunft grundsätzlich die Möglichkeit geben, Anreize für die Abgabe echter Bewertungen zu setzen. Ansatzpunkte dafür wären insbesondere eine differenzierte Behandlung seriöser Bewertungssysteme und die Öffnung großer Plattformen wie Amazon für externe Bewertungsvermittler (s. Kapitel F.2.).

E.4. Portale

Dreh- und Angelpunkt von Nutzerbewertungen im Internet sind Portale, die als Bindeglied zwischen Anbieter und Verbraucher fungieren. Um Bewertungen anzeigen zu können, die dem Nutzer in der Fülle der Angebote Orientierung bieten, betreiben diese Portale regelmäßig ein Nutzerbewertungssystem. Die unterschiedlichen Funktionsweisen der Bewertungssysteme von Portalen sind in Kapitel D. auf Basis der Ermittlungsergebnisse der Sektoruntersuchung bereits ausführlich dargestellt worden. Grundsätzlich haben Portale ein Interesse daran, möglichst viele, glaubwürdige und informative Nutzerbewertungen auf ihrer Seite darzustellen. Die konkrete Bedeutung der Nutzerbewertungen und die konkrete Interessenlage können dabei aber je nach Geschäftsmodell des Portals unterschiedlich ausfallen.

Bei Konflikten über die dargestellten Bewertungen zwischen bewerteten Unternehmen und den Verfassern von Bewertungen bewegt sich die Rolle der Portale in dem Spannungsfeld zwischen der eines neutralen Außenstehenden, der lediglich eine Plattform zur Verfügung stellt, auf der Inhalte Dritter dargestellt werden, und der eines Beteiligten, der durch die Ausgestaltung seines Bewertungssystems einen aktiven Beitrag zu den auf seiner Seite veröffentlichten Bewertungen leistet. Die Rechtsprechung der letzten Jahre hat teilweise dazu geführt, dass Portale systemseitig im Vorfeld möglichst wenige Bewertungen herausfiltern oder modifizieren, sondern sich auf eine nachträgliche Einzelfallprüfung bei Beschwerden konzentrieren (siehe Kapitel D.2.4.). Von politischer und gesetzgeberischer Seite mehrten sich derzeit jedoch Vorstöße, dass Portale über den konkreten Einzelfall hinaus mehr Verantwortung für Bewertungen übernehmen und bezüglich ihrer Filtermechanismen transparenter werden.

E.4.1. Einfluss von Bewertungen auf den Umsatz der Portale

Nutzerbewertungssysteme werden von einer ganzen Reihe von Portalen eingesetzt, die ansonsten unterschiedliche Geschäftsmodelle verfolgen. Entsprechend verschieden kann damit auch die Bedeutung sein, die Nutzerbewertungen für ein Portal bzw. dessen Umsätze haben. Diese Vielfalt spiegelt sich auch in den Ergebnissen der Befragung wider.

Im Bereich Produktbewertungen waren die Befragten überwiegend Händler oder Handelsplattformen bzw. eine Kombination von beidem. Zu den Händlern gehören Anbieter, die vorwiegend stationär tätig sind und inzwischen auch online vertreiben wie die Drogeriemarktkette dm, und Anbieter, die ganz oder überwiegend das Internet als Vertriebskanal nutzen wie die Online-Apotheke DocMorris. Zu den Handelsplattformen gehören reine Plattformen, die nicht selbst Produkte verkaufen wie Ebay, oder hybride Portale wie Amazon, auf deren Plattform sowohl sie selbst wie auch Dritthändler tätig sind. Bereits bei der Auswahl der zu befragenden Portale hat sich gezeigt, dass inzwischen alle großen Händler und Handelsplattformen Nutzerbewertungen anzeigen. Einige haben allerdings berichtet, dass sie ihr Bewertungssystem erst vor relativ kurzer Zeit eingerichtet hätten. Dies zeigt, dass Nutzerbewertungen im Onlinebereich in den letzten Jahren an Bedeutung gewonnen haben und inzwischen vom Nutzer vorausgesetzt werden; einer der Befragten hat Nutzerbewertungen in diesem Zusammenhang als „Marktstandard“ bezeichnet.

Die Aussagen der bei der Portalbefragung befragten Händler belegen, dass Bewertungen von den Nutzern tatsächlich als Entscheidungshilfe verwendet werden: Ein Händler berichtete, dass in einem Viertel aller Fälle neben den auf der Produktdetailseite angezeigten Bewertungen aktiv noch die Ansicht weiterer Bewertungen angeklickt würde. Ein anderer gab an, dass Produkte mit vielen und guten Bewertungen eine höhere Konversionsrate aufweisen, also das Anklicken auch häufiger zum Kauf führt. Die hohe Bedeutung von Bewertungen für den Verbraucher führt dazu, dass Nutzerbewertungen auch die Gesamtattraktivität eines Portals aus Verbrauchersicht erhöhen. Portale mit vielen Bewertungen verzeichnen deshalb mehr Besuche, was wiederum zu höheren Umsätzen des Portals insgesamt führt, auch wenn der Effekt von Nutzerbewertungen für ein Portal als Ganzes quantitativ geringer sein dürfte als der für ein einzelnes Produkt, dessen Bewertungen einen unmittelbaren Wettbewerbsfaktor im Vergleich zu einem Konkurrenzprodukt darstellen.

Insgesamt haben die befragten Portale im Bereich Produktbewertungen angegeben, dass ihre Umsätze ohne die Darstellung von Bewertungen sinken würden, aber nur in geringem Umfang. Nutzerbewertungen sind für die meisten Portale im Bereich Produktbewertungen demnach zwar wichtig, aber nicht entscheidend für ihr Geschäftsmodell, wie die Antworten auf die entsprechende Frage gezeigt haben (s. nachfolgende Abbildung, blaue Säulen). Ein Drittel der Portale bezeichnet die Bedeutung von Produktbewertungen als „hoch“, aber kein einziges als „sehr hoch“.

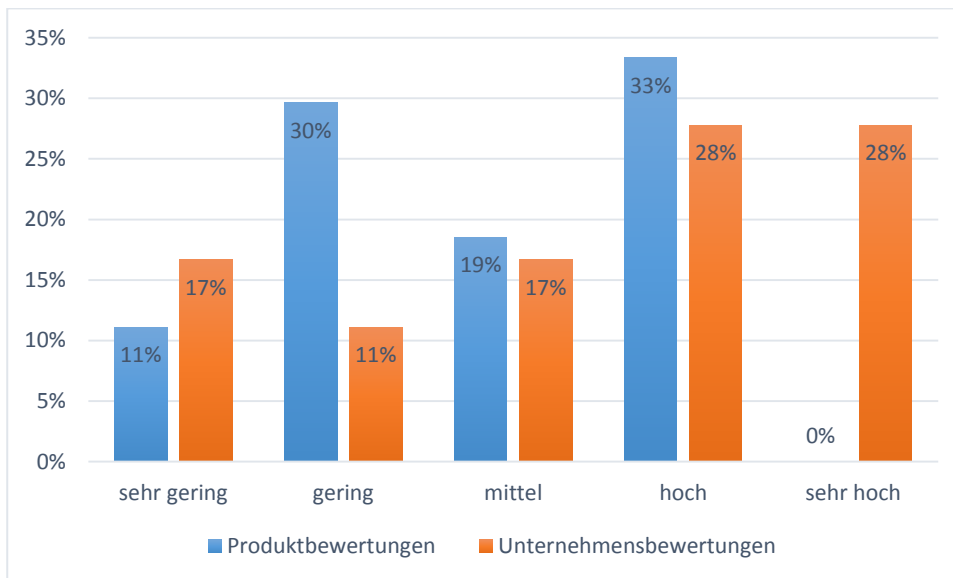


Abbildung 14: Bedeutung von Nutzerbewertungen für das Geschäftsmodell des Portals

Im Bereich Unternehmensbewertungen wurden bezogen auf ihr Geschäftsmodell sehr unterschiedliche Portale befragt. Zu den befragten Portalen gehörten Handelsplattformen wie Amazon und Ebay, die neben den Produkt- auch Händlerbewertungen anzeigen. Befragt wurden außerdem Vergleichsportale, auf denen der Kunde eine große Bandbreite von Anbietern vergleichen und deren Leistungen dann über das Portal teilweise auch bestellen kann. Manche dieser Anbieter bieten ein breites Portfolio von Vergleichen an wie z. B. Check24 oder Dasoertliche.de, andere sind auf bestimmte Branchen spezialisiert wie Airbnb, Booking oder Jameda. Je nach Geschäftsmodell wird dabei die Bedeutung der Unternehmensbewertungen sehr unterschiedlich eingeschätzt (s. obige Abbildung, orange Säulen).

Zu den Portalen, die die Bedeutung von Unternehmensbewertungen als gering oder sehr gering einschätzen, gehören die Handelsplattformen, die sowohl Produkt- als auch Unternehmensbewertungen anzeigen wie z. B. Amazon und Ebay. Zwar waren nach den Ermittlungen die Bewertungen der Händler durch andere Nutzer auf diesen Plattformen zu Beginn des Internethandels von großer Bedeutung, da sie eine Möglichkeit darstellten, Bedenken von Nutzern zu zerstreuen, bei einem völlig unbekanntem Anbieter einen Kauf zu tätigen. Inzwischen haben Nutzerbewertungen als vertrauensbildende Maßnahme in Bezug auf Händler jedoch keine so hohe Bedeutung mehr. Teilweise überträgt sich das Vertrauen der Nutzer in eine Plattform vermutlich auf deren Händler, indem sie „bei Amazon“ bestellen, auch wenn ein Dritthändler der eigentliche Vertragspartner ist. Teilweise greifen Händler aber auch auf die Gütesiegel von Dienstleistern wie Trusted Shops oder eKomi zurück, um ihre Vertrauenswürdigkeit nachzuweisen. Und schließlich haben sich die Verbraucher zunehmend daran gewöhnt, bei ihnen unbekanntem Internethändlern einzukaufen. Entsprechend haben heute Händlerbewertungen auf Handelsplattformen eine wesentlich geringere Bedeutung als Produktbewertungen und stehen deutlich weniger im Fokus der Nutzer. Der direkte Effekt von Unternehmensbewertungen auf die Umsätze derartiger Plattformen als Ganzes wird daher auch eher als gering eingeschätzt. Für das interne Monitoring der Händler durch die Portale stellen die Unternehmensbewertungen allerdings neben anderen Kriterien nach wie vor eine Möglichkeit dar, die Qualität und Zuverlässigkeit eines Händlers einzuschätzen.

Weitere Adressaten der Portalbefragung im Bereich Unternehmensbewertungen waren Plattformen, die für unterschiedliche Branchen mit vielen kleinen Anbietern deren Angebote gebündelt darstellen und vergleichbar machen. Diese Plattformen werden als Vergleichsportale, Branchensuchportale oder Suchmaschinen bezeichnet. Das damit verbundene Geschäftsmodell, das inzwischen weit verbreitet ist, besteht darin, die entsprechenden Angebote gegen eine Provision zu vermitteln oder sich die Weiterleitung über ein Pay-per-Click-Modell o. Ä. vergüten zu lassen. Zu den in diesem Bereich befragten Plattformen gehören z. B. Check24 und Expedia. Insbesondere auf Plattformen, die zumindest teilweise auch im C2C-Bereich tätig sind, können gegenseitige Bewertungen von beiden Marktseiten abgegeben werden: Bei Airbnb bewerten z. B. Gäste die Vermieter der Unterkünfte und die Vermieter ihre Gäste. Welche Rolle die auf diesen Seiten dargestellten Bewertungen für die Nutzer spielen, hängt davon ab, ob es sich um homogene und standardisierte Leistungen oder um stark differenzierte und ggfs. emotionale Erfahrungsgüter handelt. So sind Bewertungen bei der Suche nach einer Flugreise, einem Versicherungs- oder Kreditvertrag für Nutzer von untergeordneter Bedeutung, während sie bei der Suche nach einem Hotel oder Restaurant eine große Rolle spielen.⁸⁶ Entsprechend unterschiedlich wurden bei der Befragung dieser Plattformen die Auswirkungen von Nutzerbewertungen auf das Geschäftsmodell und die Umsätze der einzelnen Portale eingeschätzt.

Alternativ zum Geschäftsmodell der Vermittlung gibt es reine Bewertungsportale, bei denen keine Vermittlung stattfindet. Jameda, Tripadvisor und Yelp haben zumindest bis vor einigen Jahren keine Vermittlungsdienstleistungen angeboten, sondern finanzierten sich ausschließlich über Listungsentgelte und Anzeigen. Diese Bewertungsportale sind in Branchen tätig, in denen die Erfahrungen anderer Nutzer für Verbraucher eine besonders hohe Bedeutung haben, da es sich um sehr heterogene oder emotionale Leistungen wie z. B. von Restaurants, Friseuren, Anwälten oder Ärzten handelt. Bewertungen sind für diese Portale daher Kernstück ihres Geschäftsmodells und haben dementsprechend einen hohen Einfluss auf den Umsatz des Portals und eine sehr hohe Bedeutung für dessen Geschäftstätigkeit.

Insgesamt hat die Befragung bestätigt, dass Bewertungen und damit Bewertungssysteme nicht nur für Anbieter, sondern auch für Portale wichtig sind, teilweise als eines von mehreren Marketinginstrumenten, das aber zunehmend unverzichtbar erscheint, teilweise als Kernstück der eigenen Geschäftstätigkeit. Damit dieser Effekt von Bewertungen funktionieren kann, muss allerdings eine bestimmte Mindestmenge an Bewertungen vorliegen: Marktteilnehmer haben berichtet, dass sie in einem System erst etwa ein Jahr lang Bewertungen sammeln, aber noch nicht anzeigen, um dann eine ausreichende Masse anzeigen zu können. Außerdem müssen die Bewertungen auf einem Portal aus Sicht der Nutzer glaubwürdig erscheinen, was nicht immer der Fall zu sein scheint.

⁸⁶ Bundeskartellamt, Sektoruntersuchung Vergleichsportale, S. 113 ff., April 2019, https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Sektoruntersuchungen/Sektoruntersuchung_Vergleichsportale_Bericht.html (abgerufen am 01.10.2020).

E.4.2. Politische und gesetzgeberische Initiativen in Bezug auf Portale

Mit der zunehmenden Bedeutung von Bewertungen wächst in den letzten Jahren auch die Kritik an möglichen Missständen bei Bewertungen. Der Fokus liegt dabei fast ausschließlich auf einem möglichen Missbrauch der Systeme durch die Abgabe nicht-authentischer, manipulierter bzw. incentivierter Bewertungen, hingegen selten oder nie auf fehlender Repräsentativität der Bewertungen. Eine ganze Reihe von Institutionen und der Gesetzgeber haben sich in den letzten Jahren damit beschäftigt, welche Aufgaben den Portalen, die ein Bewertungssystem betreiben, zukommen, und welche Verantwortung sie für mögliche Missstände tragen.

Derzeit sprechen die auch als „Plattformprivileg“ bezeichneten Normen des § 7 Abs. 2 und § 10 TMG sog. „Diensteanbieter“ unter bestimmten Bedingungen von der Verantwortung für „fremde Informationen“ frei, sofern sie „keine Kenntnis von der rechtswidrigen Handlung oder der Information haben und ihnen im Falle von Schadensersatzansprüchen auch keine Tatsachen oder Umstände bekannt sind, aus denen die rechtswidrige Handlung oder die Information offensichtlich wird“. Darüber hinaus schränken diese Vorschriften die Überwachungspflichten der Diensteanbieter stark ein. Die höchstrichterliche Rechtsprechung legt das Haftungsprivileg bislang eher plattformfreundlich aus. So zieht die Leitentscheidung „Hotelbewertungsportal“ die Grenze zumutbarer Überwachungspflichten recht eng.⁸⁷ Die Voraussetzungen für ein Zu-Eigen-Machen von Nutzerbewertungen, wodurch diese zu „eigenen Informationen“ werden können, werden ebenfalls im Sinne der Plattformen relativ großzügig ausgelegt: So sei die statistische Auswertung der Bewertungen zu Durchschnittswerten und einer Weiterempfehlungsrate nicht mit einer inhaltlich-redaktionellen Kontrolle vergleichbar und führe demzufolge auch nicht dazu, dass das Portal sich den Inhalt der Bewertungen zu Eigen mache.⁸⁸

Die Tatsache, dass ein großer Teil der Portale derzeit weniger proaktive Überprüfungen von Bewertungen auf mögliche Rechtsverstöße vornehmen, als es ausweislich der Vorgehensweise des anderen Teils der Portale technisch möglich wäre, wird von Politik und Verbraucherschützern allerdings zunehmend als Problem wahrgenommen. Bei den Diskussionen geht es teilweise isoliert um Nutzerbewertungen, teilweise sind Nutzerbewertungen aber auch nur ein Aspekt in einer breiteren Diskussion über die generelle Verantwortung von Plattformen für Inhalte, die auf ihnen dargestellt werden. Gemeinsam ist allen Initiativen, dass sie die Verantwortung für Probleme bei Nutzerbewertungen bei den Portalen sehen, die die Bewertungssysteme betreiben und die nach derzeitigem Recht und der bis heute dazu entwickelten Rechtsprechung bisher nur sehr eingeschränkt für die auf ihrer Seite veröffentlichten Inhalte haften. Die Portale sind daher auch die Hauptadressaten der verschiedenen Aktivitäten.

Zu den Initiativen, die sich isoliert mit Nutzerbewertungen beschäftigen, gehört zunächst das Schwerpunktthema des ICPEN (International Consumer Protection and Enforcement Network), „Online Reviews and Endorsement“, im Jahr 2015/16. Als Ergebnis der Arbeit im ICPEN sind u. a. mehrere Richtlinien zum Umgang

⁸⁷ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 41 f. – *Hotelbewertungsportal*.

⁸⁸ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 28 – *Hotelbewertungsportal*.

verschiedener Marktteilnehmer mit Nutzerbewertungen erarbeitet worden.⁸⁹ Im Mai 2017 hat die Weltorganisation des Tourismus der Vereinten Nationen (UNWTO) „Recommendations on the Responsible Use of Ratings and Reviews on Digital Platforms“ veröffentlicht.⁹⁰ Im Jahr 2018 wurde vom Technischen Komitee ISO/TC 290 „Online reputation“, dessen Sekretariat bei der französischen Stelle für Normung, AFNOR (Association française de normalisation), angesiedelt ist, eine ISO Norm für Online-Bewertungssysteme erarbeitet, die inzwischen auch in Deutschland als DIN-Norm verfügbar ist.⁹¹ Nach den Ermittlungen im Rahmen der Sektoruntersuchung haben sich in Deutschland allerdings bisher kaum Unternehmen nach dieser Norm zertifizieren lassen. Der vzbv hat sich ebenfalls verschiedentlich mit Nutzerbewertungen verfasst. So wurde 2018 eine Studie veröffentlicht, die schwerpunktmäßig untersucht hat, welche Maßnahmen Portale bei der Filterung ihrer Bewertungen vornehmen.⁹² Telefoninterviews mit mehreren Portalen im Bereich Elektronikprodukte haben dabei große Unterschiede in der Tiefe der Prüfung ergeben, die zu einer Kategorisierung der Portale in drei Gruppen geführt hat: die Gewissenhaften, die Ambitionierten und die Zurückhaltenden. Die große Heterogenität beim Umgang der Portale mit Nutzerbewertungen hat sich in der Befragung des Bundeskartellamts bestätigt.

Etwas umfassender mit dem Online-Handel allgemein, dabei aber auch mit Nutzerbewertungen, beschäftigt sich ein vom vzbv in Auftrag gegebenes Gutachten zum Verbraucherschutz im Onlinehandel, in dem u. a. geprüft wird, welche „Designpflichten“ ein Portal nach geltender Rechtslage zu erfüllen hat.⁹³ Auch die Gesetzgeber in einzelnen europäischen Ländern und bei der europäischen Kommission beschäftigten sich in den vergangenen Jahren mit den Herausforderungen, die Online-Handel und Online-Plattformen mit sich bringen. In Frankreich wurde 2016 ein „Gesetz für eine digitale Republik“ verabschiedet,⁹⁴ das u. a. eine ergänzende Vorschrift zu Nutzerbewertungen in den Code de la consommation⁹⁵ aufnimmt: Portale, die ein Nutzerbewertungssystem betreiben, sind demnach dazu verpflichtet, anzugeben, ob die Bewertungen einer Kontrolle unterworfen werden, und wenn ja, in welcher Form. Bewertungen sind zudem mit dem Datum ihrer Abgabe zu versehen. Wird eine Bewertung nicht angezeigt, ist der Verfasser darüber zu informieren und die Ablehnung zu begründen. Außerdem müssen die Bewerteten nach dieser Vorschrift eine kostenlose

⁸⁹ International Competition Network, Online Reviews and Endorsements, ICPEN Guidelines, 2016 <https://www.icpen.org/initiatives> (abgerufen am 01.10.2020).

⁹⁰ Siehe <https://www.unwto.org/archive/europe/press-release/2017-05-01/unwto-international-congress-ethics-presents-recommendations-responsible-us> (abgerufen am 01.10.2020).

⁹¹ DIN ISO 20488, Online-Kundenbewertungen – Grundsätze und Anforderungen für die Erhebung, Moderation und Veröffentlichung, 2019.

⁹² VZBV, Fälschungen bei Bewertungen – Bekämpfen Online-Portale sie wirksam?, April 2018.

⁹³ Busch, Effektiver Verbraucherschutz im Online-Handel: Verantwortung und Haftung von Internetplattformen, Rechtsgutachten im Auftrag der Verbraucherzentrale Bundesverband e. V., Abschnitt 5 Kundenbewertungen, S. 48 ff., 2019.

⁹⁴ LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique, <https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000033202746> (abgerufen am 01.10.2020).

⁹⁵ Article L111-7-1, Code de la Consommation, Fassung vom 30. Januar 2020, <https://www.legifrance.gouv.fr/affich-Code.do?cidTexte=LEGITEXT000006069565> (abgerufen am 01.10.2020).

Möglichkeit haben, die Authentizität einer Bewertung begründet anzuzweifeln. Die Details der Vorschrift regelt eine Verordnung.⁹⁶

Die EU-Kommission hat in einem breit angelegten „New Deal for Consumers“ verschiedene Richtlinien, die Verbraucherrechte betreffen, ergänzt. Dabei sind einige Vorschriften zu Nutzerbewertungen in die UGP-Richtlinie aufgenommen worden, die derzeit vom Bundesministerium der Justiz und für Verbraucherschutz in deutsches Recht umgesetzt werden. So wird im neuen Art. 7 Abs. 6 UGP-Richtlinie („Irreführende Unterlassungen“) klargestellt, dass Informationen darüber, ob und wie ein Gewerbetreibender sicherstellt, dass die veröffentlichten Bewertungen von Verbrauchern stammen, die die Produkte tatsächlich verwendet oder erworben haben, als wesentlich gelten. Dies hat Auswirkungen auf die Anwendbarkeit z. B. der Generalklausel des § 3 Abs. 2 UWG („Verbot unlauterer geschäftlicher Handlungen“), nach der eine wesentliche Beeinflussung des wirtschaftlichen Verbraucherverhaltens Voraussetzung für die Unlauterkeit einer geschäftlichen Handlung ist. Anhang I der UGP-Richtlinie („Geschäftspraktiken, die unter allen Umständen als unlauter gelten“), wurde ebenfalls ergänzt. Zu den ohne Weiteres irreführenden Geschäftspraktiken gehört nun auch die „Behauptung, dass Bewertungen eines Produkts von Verbrauchern stammen, die das Produkt tatsächlich verwendet oder erworben haben, ohne dass angemessene und verhältnismäßige Schritte unternommen wurden, um zu prüfen, ob die Bewertungen wirklich von solchen Verbrauchern stammen“ (Nr. 23b). Gleiches gilt für „Die Abgabe gefälschter Bewertungen oder Empfehlungen von Verbrauchern bzw. die Erteilung des Auftrags an andere juristische oder natürliche Personen, gefälschte Bewertungen oder Empfehlungen von Verbrauchern abzugeben, sowie die falsche Darstellung von Verbraucherbewertungen oder Empfehlungen in sozialen Medien zu Zwecken der Verkaufsförderung“ (Nr. 23c). Die ab Juli 2020 für das Verhältnis zwischen Plattformen und Anbieter geltende P2B-Verordnung⁹⁷ beinhaltet keine expliziten Regelungen zum Thema Nutzerbewertungen.

Parallel zu diesen Initiativen sind Wettbewerbs- und Verbraucherschutzbehörden in verschiedenen Ländern mit konkreten Verfahren gegen Verbraucherschutzverstöße bei Nutzerbewertungen vorgegangen (s. Kapitel B.).

E.4.3. Wesentliche Erkenntnisse

Bewertungen sind insbesondere für Verbraucher und Anbieter von großer Bedeutung. Nutzerbewertungen sind in vielen Branchen ein Standard, der zunehmend von den Verbrauchern erwartet wird. Je nach Geschäftsmodell und Art der Bewertungen sind die konkreten Interessenlagen der einzelnen Portale jedoch unterschiedlich. Insbesondere bei stark differenzierten und emotionalen Leistungen können Nutzerbewertungen das Kernstück eines Portals darstellen und sind damit für dessen Geschäftstätigkeit unverzichtbar.

⁹⁶ Décret n° 2017-1436 du 29 septembre 2017 relatif aux obligations d'information relatives aux avis en ligne de consommateurs, <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000035720935&fastPos=1&fastReqId=504132173&categorieLien=id&oldAction=rechTexte> (abgerufen am 01.10.2020).

⁹⁷ Verordnung (EU) 2019/1150 des Europäischen Parlaments und des Rates v. 20.06.2019 zur Förderung von Fairness und Transparenz für gewerbliche Nutzer von Online-Vermittlungsdiensten, Abl. EU Nr. L 186 v. 11.07.2019, S. 57.

Umso wichtiger ist es für die Portale, eine hinreichende Zahl von repräsentativen und glaubwürdigen Nutzerbewertungen darstellen und gegen potenzielle Beeinträchtigungen dieses wichtigen Erfolgsfaktors effektiv vorgehen zu können. Diese teilweise entgegenlaufenden Interessen und die teilweise unklare bzw. dynamische Rechtslage bilden die Rahmenbedingungen für das Verhalten der Portale in Bezug auf Nutzerbewertungen.

Das Problem nicht-authentischer, manipulierter und incentivierter Bewertungen wird auch auf der politischen Ebene zunehmend wahrgenommen, teilweise auch im größeren Zusammenhang der Regelungen zur Verantwortlichkeit von Portalen insgesamt für die auf ihnen dargestellten Inhalte. Im Fokus stehen dabei eine umfassendere Aufklärung der Verbraucher über die existierenden Kontrollmechanismen eines Portals und die Bekämpfung von „gefälschten“ Bewertungen. Die mangelnde Repräsentativität von Nutzerbewertungen, die u. a. durch Prozesse wie Notice-and-take-down noch verstärkt wird, spielt in der politischen Diskussion hingegen kaum eine Rolle.

F. Problemdifferenzierung und Lösungsansätze

Die Sektoruntersuchung hat gezeigt, dass bei Nutzerbewertungen systembedingt Probleme bestehen, die verhindern, dass diese wichtige Möglichkeit des Informationsaustausches zwischen Verbrauchern optimal stattfinden kann: Insgesamt gibt es zu wenig Bewertungen, weil zu wenig Verbraucher bereit sind, kostenlos Bewertungen abzugeben. Gleichzeitig sind aber gerade Anbieter, die erst wenige Bewertungen haben, darauf angewiesen. Um für mehr Bewertungen zu sorgen und gleichzeitig die Beschaffung von Bewertungen über unseriöse Kanäle einzudämmen, müssen neue Wege beschritten werden. Verbraucherrechtskonform lassen sich mehr Bewertungen z. B. durch den transparenten und manipulationsfreien Einsatz von Belohnungen oder Produkttests generieren. Dies geschieht bisher viel zu selten.

Tatsächlich wählen einige Anbieter derzeit viel zu häufig andere, nicht wünschenswerte Wege, um die Zahl an (positiven) Bewertungen für ihr Angebot zu erhöhen, was zu unterschiedlichen Missständen führt: Incentivierungen werden häufig nicht kenntlich gemacht. Es gibt zu wenige Möglichkeiten, Produkttester einzusetzen und dies entsprechend transparent zu kennzeichnen. Der Anteil nicht-authentischer oder manipulierter Bewertungen ist zu hoch. Teilweise werden zu wenige negative Bewertungen angezeigt. Durchschnittsangaben sind in einigen Fällen missverständlich. Und insgesamt ist die Basis derjenigen Verbraucher, die Bewertungen abgeben, sehr klein und vermutlich häufig nicht repräsentativ. Portale hätten nach vorläufiger Einschätzung des Bundeskartellamts in fast allen Bereichen die Möglichkeit, mehr gegen Verbraucherrechtsverstöße zu unternehmen und könnten dazu nach geltendem Recht auch verpflichtet werden. Zu einzelnen Problemkonstellationen existieren auch bereits Gerichtsentscheidungen. Nach Auffassung des Bundeskartellamtes ist eine pauschale Verurteilung von „Fake-Bewertungen“ in diesem Zusammenhang allerdings wenig zielführend, da das Thema differenziert und zudem im Spannungsfeld der geschilderten verschiedenen Interessenlagen zu betrachten ist.

Die folgende Abbildung gibt zunächst einen Überblick über die verschiedenen Arten, wie Bewertungen zustande kommen können.

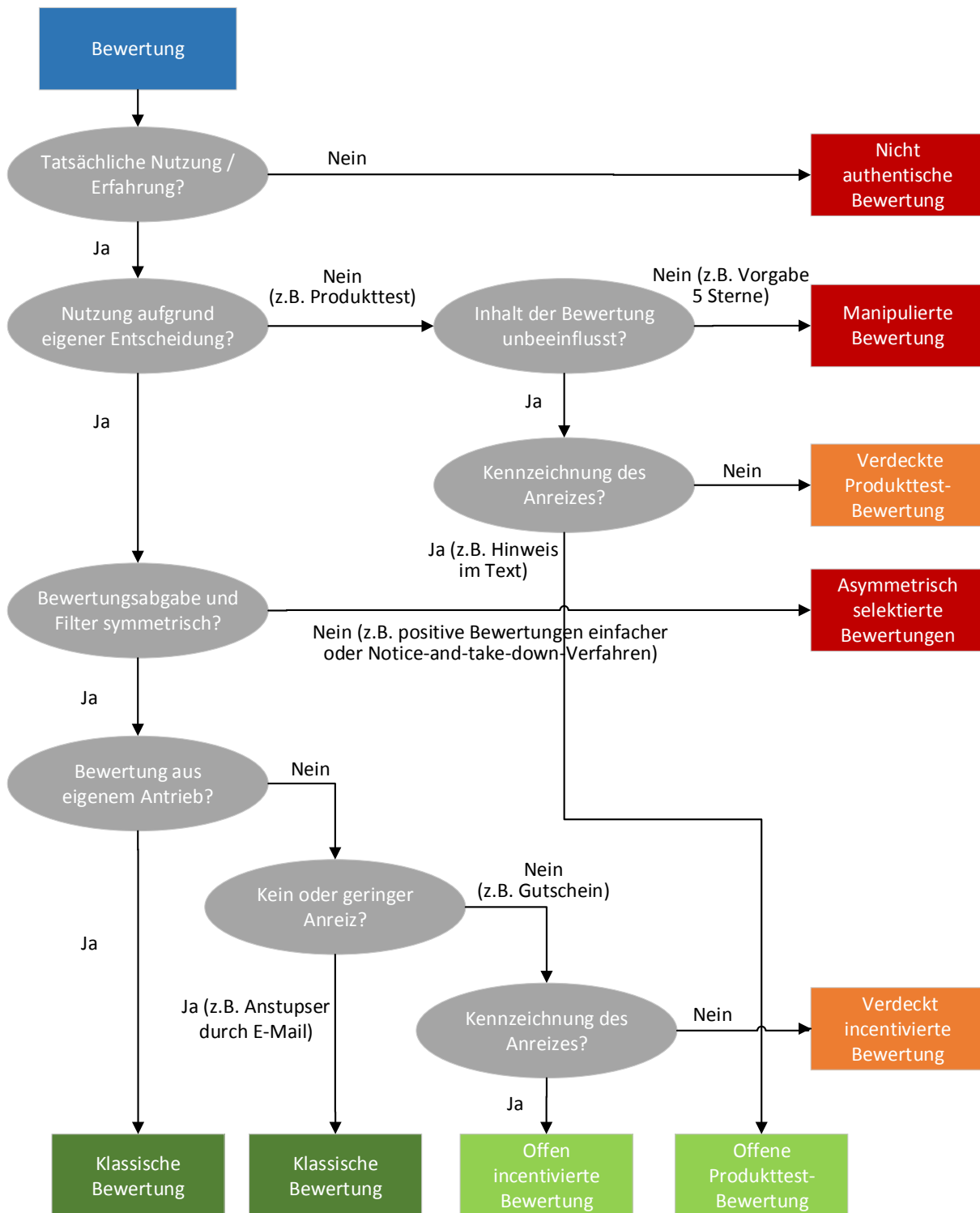


Abbildung 15: Klassifizierung von Bewertungen in Bewertungssystemen

Eine **klassische Bewertung** kommt zustande, wenn der Verfasser das Item aufgrund eigener Entscheidung selbst genutzt hat und entweder aus eigenem Antrieb entscheidet, eine Bewertung abzugeben, oder lediglich durch einen Anstupser (Erinnerung, symbolischer Euro etc.) dazu motiviert wird, eine Bewertung abzugeben, und wenn außerdem sowohl die Bewertungsmöglichkeit als auch die anschließende Filterung symmetrisch sind. Klassische Bewertungen bieten einen großen Mehrwert für Verbraucher. Sind alle Bedingungen für eine klassische Bewertung erfüllt, liegt jedoch der für die Bewertung gewährte Anreiz so hoch, dass er ggfs. die Bewertung beeinflussen könnte, handelt es sich um eine incentivierte Bewertung. Die Incentivierung kann

offen gelegt werden (**offen incentivierte Bewertung**) oder nicht (**verdeckt incentivierte Bewertung**). Darüber hinaus gibt es die Situation, in denen Nutzer das bewertete Item gar nicht unabhängig von der Bewertung für sich selbst ausgesucht haben, sondern es sich um eine Gefälligkeit oder eine Aufforderung zu einem Produkttest handelt. Nimmt der Verfasser der Bewertung dessen ungeachtet eine unbeeinflusste Bewertung vor, liegt eine Produkttest-Bewertung vor, die wiederum als solche gekennzeichnet werden kann (**offene Produkttest-Bewertung**) oder nicht (**verdeckte Produkttest-Bewertung**). Falls mit der Vorgabe bzw. Bitte, ein Item zu bewerten, auch noch Einfluss auf den Inhalt der Bewertung wie z. B. die Anzahl vergebener Sterne genommen wird, handelt es sich um eine **manipulierte Bewertung**. Bewertungen, denen keine tatsächliche Nutzung oder Erfahrung des Verfassers zugrunde liegt, werden schließlich als **nicht-authentisch** bezeichnet. Falls schließlich die Möglichkeit, eine Bewertung abzugeben, asymmetrisch ausgestaltet ist, so dass eher zufriedene Kunden bewerten, oder wenn die Filtermechanismen für Bewertungen den Effekt haben, dass negative Bewertungen eher gar nicht dargestellt oder nachträglich entfernt werden, handelt es sich um **asymmetrisch selektierte Bewertungen**. In diesem Fall spiegelt jede einzelne Bewertung zwar für sich genommen die tatsächliche Erfahrung des Verfassers, da jedoch negative Erfahrungen mit größerer Wahrscheinlichkeit nicht dargestellt werden, ist das Gesamtbild, das Leser von Bewertungen erhalten, positiv verzerrt.

Alle in der Abbildung rot bzw. orange markierten Arten von Bewertungen sind auf Portalen zu finden und stellen grundsätzlich einen Verstoß gegen das Lauterkeitsrecht dar. Die in der Graphik hellgrün markierten Arten von Bewertungen sind dagegen, obwohl sie in der öffentlichen Diskussion teilweise ebenfalls als Fake- oder gefälschte Bewertungen bezeichnet werden, zulässig und bieten Anbietern und Portalen grundsätzlich die Möglichkeit, mehr Bewertungen zu generieren, wenn die Zahl klassischer Bewertungen zu gering ist.

F.1. Incentivierte Bewertungen

Bei incentivierten Bewertungen haben die Nutzer sich aufgrund einer eigenen Entscheidung für den Kauf eines Produktes entschieden. Sie werden aber mit einem so hohen Anreiz dazu motiviert, eine Bewertung abzugeben, dass sie nicht nur mit deutlich größerer Wahrscheinlichkeit eine Bewertung abgeben, sondern ggfs. auch ein Einfluss des Anreizes auf den Inhalt der Bewertung zu erwarten ist. Diese Incentivierung besteht typischerweise in einem Gutschein oder der Teilnahme an einem Gewinnspiel und kann entweder von dem Portal als Betreiber des Bewertungssystems selbst oder von einem auf dem Portal aktiven Anbieter für dessen Leistungen ausgelobt werden. Auch Bewertungen, die aufgrund eines Produkttests entstehen, sind incentiviert, werden jedoch als eigene Fall-Konstellation im nachfolgenden Abschnitt F.2. behandelt. Bewertungen, die über die Incentivierung hinaus inhaltlich beeinflusst, also manipuliert werden, sind Gegenstand des Abschnitts F.3.

F.1.1. Ökonomische und rechtliche Bewertung

Incentivierungen für die Abgabe einer Bewertung sind ökonomisch grundsätzlich positiv zu bewerten, da sie zu mehr Bewertungen auf breiter Basis führen und damit dazu beitragen können, die Lücke zwischen dem Angebot und der Nachfrage nach Bewertungen zu schließen. Allerdings führen derartige Anreize dazu, dass

die Items, für deren Bewertung eine Incentivierung erfolgt, häufiger und besser bewertet werden. Im Vergleich zu nicht-incentivierten Bewertungen entsteht eine Asymmetrie: Wenn für die Bewertungen der Produkte einer bestimmten Marke auf einem Portal ein Incentive vergeben wird, für alle anderen Produkte aber nicht, werden Produkte dieser Marke häufiger und besser als die anderen bewertet. Wenn ein Portal für alle Bewertungen ein Incentive vergibt, werden alle Items auf diesem Portal häufiger und besser bewertet als die Items auf anderen Portalen. Wenn Verbraucher nicht erkennen können, dass diese Unterschiede ggfs. auf den Incentive des Markenanbieters oder des Portals zurückzuführen sind, werden sie irreführt; dies ist auch bereits gerichtlich bestätigt worden.⁹⁸ Wenn die Incentivierung allerdings transparent gemacht wird, kann sie Verbrauchern nutzen, weil sie zu mehr Bewertungen und zu einem möglicherweise repräsentativeren Gesamtbild an Bewertungen führt, auch wenn die Bewertungen dann insgesamt positiver ausfallen.

Mehr Bewertungen sind zunächst hilfreich für die Verbraucher, da sie damit mehr Informationen über das betreffende Produkt erhalten. Die Tatsache, dass durch den gewährten Anreiz auch Personen eine Bewertung abgeben, die dies sonst nicht tun würden, kann zudem dazu führen, dass die Bewertungen in ihrer Summe die Vielfalt der Verbrauchererfahrungen und -meinungen besser widerspiegeln, da nicht nur eine sehr kleine Gruppe von Nutzern mit möglicherweise extremen Erfahrungen eine Bewertung abgibt (s. Kapitel E.1.2.). Aus ökonomischer Sicht ist das Mehr an Informationen zu begrüßen, weil Verbraucher sich damit vor dem Kauf besser informieren und bessere Entscheidungen treffen können.

Allerdings haben zahlreiche Studien gezeigt, dass allein die Anzahl an Bewertungen für Nutzer ein wichtiges Entscheidungskriterium ist: Häufiger bewertete Items werden häufiger ausgewählt (s. Kapitel E.2.1.). Wenn die Verbraucher über die Incentivierung nicht aufgeklärt werden (verdeckt incentivierte Bewertung), kann dies dazu führen, dass sie aus der hohen Zahl der Bewertungen fälschlicherweise Rückschlüsse auf die große Bekanntheit oder Beliebtheit des Unternehmens bzw. seiner Produkte ziehen. Eine solche Irreführung ist ökonomisch nicht sinnvoll, da sie die Verbraucher zu falschen Entscheidungen verleitet, und außerdem einen Verbraucherrechtsverstoß darstellt, wie auch das Oberlandesgericht Frankfurt/M. in einem Urteil festgestellt hat,⁹⁹ in dem es um Bewertungen für einen Anbieter von Whirlpools ging. Dabei wurden Nutzer für die Abgabe von Bewertungen, Likes u. a. auf verschiedenen sozialen Netzwerken mit der Teilnahme an einem Gewinnspiel belohnt, bei dem ein Whirlpool zu gewinnen war. Das Gericht sah in der durch das Gewinnspiel gesteigerten Anzahl an Bewertungen eine Irreführung: Da nicht explizit auf den Anreiz hingewiesen wurde, hätten Besucher der Seite die Bewertungen zumindest inhaltlich durchgehen müssen, um möglicherweise erkennen zu können, dass sie teilweise nur anlässlich des Gewinnspiels abgegeben wurden.¹⁰⁰

Eine Incentivierung führt außerdem zu positiveren Bewertungen, d. h. die Verfasser bewerten das Produkt aufgrund der Belohnung, die sie fürs Bewerten bekommen, tendenziell positiver als ohne Belohnung, selbst wenn die Incentivierung explizit unabhängig vom Inhalt der Bewertung erfolgt. Dies ist auf positive Rezipro-

⁹⁸ Vgl. OLG Frankfurt/M., Urte. v. 16.05.2019, Az. 6 U 14/19.

⁹⁹ OLG Frankfurt/M., Urte. v. 16.05.2019, Az. 6 U 14/19.

¹⁰⁰ OLG Frankfurt/M., Urte. v. 16.05.2019, Az. 6 U 14/19, = juris Rn. 27, 31.

zität zurückzuführen. Positiv reziprok handelnde Menschen geben, wenn ihnen jemand etwas anbietet, uneigennützig etwas zurück, d. h. sie revanchieren sich für uneigennütziges Handeln ihres Gegenübers. Unabhängig davon, ob es sich dabei um einen inhärenten Wesenszug oder eine erlernte soziale Norm handelt, zeigen Beobachtungen aus der realen Welt ebenso wie Laborexperimente, dass etwa die Hälfte aller Personen – je nach den Umständen etwas mehr oder weniger – sich positiv reziprok verhält.¹⁰¹ Übertragen auf die Abgabe von Bewertungen führt dies dazu, dass zumindest ein Teil der Nutzer, wenn er für die Abgabe einer Bewertung eine Belohnung bekommt, wie z. B. einen Gutschein oder die Teilnahme an einem attraktiven Gewinnspiel, eine eher positivere Bewertung abgibt, auch ohne dazu in irgendeiner Weise aufgefordert zu werden. Das dürfte zumindest dann der Fall sein, wenn der Anbieter derjenige ist, der die Belohnung vergibt. Nach Auffassung des Oberlandesgerichts Frankfurt/M. im Falle der Whirlpool-Bewertungen „liegt [...] auf der Hand, dass Bewertungen aus Anlass des Gewinnspiels eher positiv ausfallen.“ Es sei damit zwar keine „bezahlte“ Werbung im Wortsinn gegeben, gleichwohl seien die Bewertungen nicht als objektiv anzusehen, so dass die Verbraucher beim Lesen der Bewertungen irregeführt würden¹⁰² – auch hier in einer Konstellation, in der bei der Darstellung der Bewertungen nicht auf den Anreiz hingewiesen wird (verdeckt incentivierte Bewertung). Die ins Positive verzerrten Bewertungen führen indes nicht dazu, dass so entstandene Bewertungen keinerlei Informationsgehalt für Verbraucher haben. Verbraucher sind z. B. aus der Werbung daran gewohnt, tendenziöse Informationen zu verarbeiten. Dazu müssen sie allerdings darüber aufgeklärt werden, dass eine Incentivierung vorliegt, so dass nicht wie sonst bei Bewertungen von einer unbeeinflussten Meinungsäußerung auszugehen ist.

Zusammenfassend können incentivierte Bewertungen dem Verbraucher helfen, bessere Entscheidungen zu treffen, wenn sie darüber aufgeklärt werden. Fehlt diese Aufklärung – z. B. in Form einer Kennzeichnung, so ist sowohl in der insgesamt erhöhten Zahl der Bewertungen als auch in der im Durchschnitt erhöhten Sternanzahl eine ökonomisch zu kritisierende Asymmetrie als auch rechtlich eine Irreführung des Verbrauchers nach § 5 Abs. 1 UWG zu sehen.¹⁰³

F.1.2. Lösungsansatz: Bedarfsgerechter Einsatz von gekennzeichneten Anreizen

Grundsätzlich ist es unter Beachtung der lauterkeitsrechtlichen Vorgaben zulässig, Nutzern Anreize zu setzen, damit sie Bewertungen abgeben. In Situationen, in denen Items zu wenige klassische Bewertungen erhalten, sollten insbesondere Portale daher verstärkt die Möglichkeit nutzen, Verbraucher durch kleine Belohnungen zur Abgabe von Bewertungen zu motivieren. Denn eine große Anzahl an Bewertungen ist nicht nur im Interesse des Portals bzw. Anbieters, sondern auch im Interesse der Verbraucher, für die Bewertungen eine wichtige Entscheidungshilfe darstellen. Diese Anreize müssen aber, wenn sie zu starken Asymmetrien im Vergleich zu nicht-incentivierten Bewertungen führen, offen gelegt werden. Ein Verbraucher sollte dann in der

¹⁰¹ *Fehr/Gächter*, Fairness and Retaliation: The Economics of Reciprocity, *Journal of Economic Perspectives*, 2000, Band 14, S. 159-181.

¹⁰² OLG Frankfurt/M., Urt. v. 16.05.2019, Az. 6 U 14/19, = juris Rn. 27.

¹⁰³ Vgl. OLG Frankfurt/M., Urt. v. 16.05.2019, Az. 6 U 14/19, = juris Rn. 31.

Lage sein, eine entsprechend hohe Anzahl von positiven Bewertungen nicht (nur) auf die Beliebtheit des Produkts, sondern auch auf den vom Portal oder Anbieter gewährten Anreiz zurückzuführen.

Im Einzelfall mag allerdings unklar sein, wann ein Anreiz so stark verzerrend ist, dass eine kennzeichnungspflichtige Incentivierung vorliegt. Auf der sicheren Seite sind Anbieter und Portale, wenn sie jede Form des Anreizes kennzeichnen. Zu berücksichtigen sind ansonsten bei dieser Frage die Höhe des Anreizes und die davon ausgehende Selektionswirkung.

Bezogen auf die Höhe eines Anreizes könnte man im Extremfall die Ansicht vertreten, dass bereits Anstupser der Form, dass z. B. ein Nutzer per Email zur Bewertungsabgabe aufgefordert wird, eine Irreführung darstellen. Denn auch diese Anstupser können (und sollen) dazu führen, dass mehr Verbraucher eine Bewertung abgeben, so dass Nutzer denken könnten, dass es sich um ein besonders beliebtes Item handelt. Es ist auch nicht auszuschließen, dass sich selbst in dieser Konstellation einzelne Verbraucher verpflichtet fühlen, eine positive Bewertung abzugeben, um sich z. B. für eine reibungslose Lieferung eines Items zu bedanken. Allerdings sollte hier noch der positive Aspekt einer größeren Zahl von Bewertungen auf einer breiteren Basis dazu führen, dass die Verbraucher durch die Anstupser unter dem Strich eher besser informiert und nicht etwa irreführt werden.

Werden über Anstupser hinaus echte monetäre Anreize wie Gutscheine oder Gewinnspiele gewährt, entscheidet die Höhe des Anreizes über die Stärke der dadurch bewirkten Verzerrung sowohl bezüglich der Zahl der Bewertungen als auch bezüglich des Inhalts bzw. der vergebenen Sterne. Wo die Grenze liegt, ab der es zwingend notwendig ist, eine Incentivierung offen zu legen, ist bislang nicht geklärt. Das von Marktteilnehmern teilweise als Orientierung herangezogene Urteil „Zugabenbündel“ des Bundesgerichtshofs¹⁰⁴ ist nicht nur bereits recht alt, sondern betrifft vor allem einen anderen Sachverhalt. In dem Fall ging es um ein Bündel aus Warenproben (Vitamintabletten, ein Stück Seife) und geringwertigen Reklameartikeln (ein Notizblock, auf dem der Name der Apotheke aufgedruckt war), denen der Kunde in Summe einen Wert von mehr als 1 DM zuweist, und die tatsächlich einen Wert von etwa 1,75 DM hatten. Der Bundesgerichtshof hatte zu entscheiden, ob die Übergabe dieses Bündels einen Verstoß gegen § 1 Abs. 1 Satz 1 ZugabeVO darstellt. Eine verbraucherrechtliche Irreführung war nicht Gegenstand der Entscheidung, so dass eine Übertragbarkeit dieser Geringfügigkeitsschwelle auf Nutzerbewertungen offen ist. In jedem Fall kann eine Irreführung dadurch ausgeschlossen werden, dass im Umfeld der angezeigten Bewertungen bzw. der Zusammenfassung der Bewertungen gut sichtbar auf den entsprechenden Anreiz hingewiesen wird (offen incentivierte Bewertung). Der vzbv empfiehlt in seiner Stellungnahme zum Konsultationspapier eine generelle Kennzeichnungspflicht für incentivierte Bewertungen, wenn die Anreize einen monetär-materiellen Charakter haben. Anreize, die sich nicht monetär auswirken, wie etwa „Statuspunkte“ im Sinne einer Gamification, sollten von einer Kennzeichnungspflicht nicht umfasst sein.

Für die Selektionswirkung eines Anreizes ist vor allem wichtig, ob der Anreiz nur für einzelne Items bzw. einzelne Anbieter auf einem Portal gewährt wird oder flächendeckend für alle. Wenn ein Anreiz selektiv, d. h.

¹⁰⁴ BGH, Urt. v. 04. 07.2002, Az. I ZR 38/00, – *Zugabenbündel*.

nur für ein Item einer bestimmten Kategorie, gewährt wird, kann schon bei einem Anreiz in sehr geringer Höhe eine Irreführung vorliegen, weil Verbraucher wohl in der Regel davon ausgehen, dass die nebeneinander präsentierten Bewertungen für verschiedene Produkte auf dem gleichen Portal unter den gleichen Bedingungen zustande gekommen sind. In diesem Fall wäre ein deutlicher Hinweis bei dem betreffenden Produkt bzw. Item erforderlich. Wenn dagegen ein Portal flächendeckend für alle Items eines Bewertungssystems Anreize zur Bewertung setzt, entsteht die Verzerrung nur in Bezug auf Bewertungen auf einem anderen Portal. Hier dürften Verbraucher eher davon ausgehen, dass Bewertungen unter leicht unterschiedlichen Voraussetzungen entstanden sein könnten. Das Gewähren eines geringen allgemeinen Anreizes auf dem einen Portal führt dann noch nicht zu einer Irreführung im Vergleich zu den Bewertungen auf einem anderen und wäre daher noch nicht kennzeichnungspflichtig. Eine Kennzeichnung von Anstupsern und Anreizen auf Portalen sollte insbesondere dann nicht notwendig sein, wenn sie branchenüblich sind, weil die Nutzer dann damit vertraut sind, z. B. weil sie selbst auch schon derartige Aufforderungen zur Abgabe einer Bewertung erhalten haben.

F.2. Produkttest-Bewertungen

Klassische und incentivierte Bewertungen werden von Nutzern für Items abgegeben, für deren Nutzung sie sich grundsätzlich unabhängig von der Bewertungsmöglichkeit entschieden haben. Wenn dagegen jemand ein Item vorrangig nutzt, damit er dafür eine Bewertung erstellt, handelt es sich um einen Produkttest. Die zu bewertende Ware oder Dienstleistung wird dem Nutzer dabei in der Regel deutlich reduziert oder sogar kostenlos überlassen, in Einzelfällen erhält er sogar darüber hinaus noch ein Entgelt für die Bewertung. Produkttests liegen auch dann vor, wenn die Nutzer sich ihre zu bewertenden Produkte frei aus einer Menge von unterschiedlichen zum Produkttest angebotenen Produkten aussuchen. Denn die Entscheidung für ein Produkt bleibt auf die zur Bewertung angebotenen Produkte beschränkt und wird außerdem dadurch beeinflusst, dass der Nutzer das Produkt kostenlos oder deutlich vergünstigt erhält. Sofern im Rahmen von Produkttests eine inhaltliche Beeinflussung der Bewertungen erfolgt, handelt es sich um manipulierte Bewertungen, die Gegenstand des nachfolgenden Abschnitts F.3. sind.

F.2.1. Ökonomische und rechtliche Bewertung

Produkttest-Bewertungen sind aus ökonomischer Sicht grundsätzlich sinnvoll: Viele Verbraucher interessieren sich für das *Lesen* von Bewertungen und nutzen Bewertungen zur Entscheidungsfindung. Das *Verfassen* von Bewertungen findet allerdings nur sehr eingeschränkt statt. Da die Formulierung ausführlicher, aussagekräftiger Bewertungen mit einem nicht unerheblichen Aufwand verbunden ist, sind nur wenige Nutzer in bestimmten Situationen dazu ohne Gegenleistung bereit. Produkttests stellen – ähnlich wie die oben behandelten Incentivierungen – eine Form der Bezahlung von Bewertungsverfassern dar und können dazu beitragen, die Lücke zwischen Angebot und Nachfrage bei Bewertungen zu schließen. Wie bei incentivierten Bewertungen besteht allerdings auch hier eine grundsätzliche Tendenz, dass Items für die Produkttests durchgeführt werden, häufiger und positiver bewertet werden (siehe Abschnitt F.1.).

Auch aus rechtlicher Sicht sind Produkttest-Bewertungen grundsätzlich nicht zu beanstanden. Nur Produkttest-Bewertungen, die nicht als solche gekennzeichnet sind, lösen nach der Rechtsprechung einen Verstoß jedenfalls gegen § 5 Abs. 1 Nr. 1 bzw. § 5a Abs. 6 UWG aus: Nicht gekennzeichnete Produkttest-Bewertungen sind irreführend nach § 5 Abs. 1 Nr. 1, weil sie zur Täuschung geeignete Angaben über die Beschaffenheit sowie ggfs. weitere Merkmale einer Ware oder Dienstleistung enthalten. Diese Angaben sind geeignet, den Verbraucher zu einer geschäftlichen Entscheidung zu veranlassen, die er andernfalls nicht getroffen hätte. Denn wie zahlreiche Studien zeigen, haben Nutzerbewertungen einen Einfluss auf das Kaufverhalten. Häufiger und besser bewertete Items werden öfter ausgewählt. Nicht gekennzeichnete Produkttests können außerdem als eine verdeckte Werbung nach § 5a Abs. 6 UWG angesehen werden. Denn der kommerzielle Zweck einer Produkttest-Bewertung wird nicht kenntlich gemacht, mit der Folge, dass Verbraucher zu einer geschäftlichen Entscheidung veranlasst werden, die sie andernfalls nicht getroffen hätten, auch hier aufgrund des Einflusses von Nutzerbewertungen auf das Kaufverhalten.

Zu verdeckten Produkttests gibt es bereits Rechtsprechung. Kläger sind in der Regel die Portale, auf denen die Bewertungen veröffentlicht werden. Die überwiegende Mehrzahl der Verfahren richtet sich gegen Bewertungsvermittler. Verfahren gegen Anbieter bilden dagegen eher die Ausnahme.¹⁰⁵ Gegen die Rezensenten selbst sind nach den Ermittlungen des Bundeskartellamtes bislang keine Verfahren angestrengt worden, obwohl denkbar wäre, dass auch diese als verantwortlich für den jeweiligen Verstoß und damit als Täter anzusehen sind.¹⁰⁶ Häufig handelt es sich bei den Gerichtsentscheidungen zu verdeckten Produkttests um nicht veröffentlichte einstweilige Verfügungen. Nur in wenigen Fällen enthalten die Urteile nähere Ausführungen zum Verbraucherrechtsverstoß durch beauftragte Bewertungen. Teilweise wird dabei der Schwerpunkt der Begründung darauf gelegt, dass die Inhalte der Bewertung und die Sternebewertung vorgegeben werden oder sogar den im jeweiligen Fall betrachteten Bewertungen keine echten Verbrauchererfahrungen zugrunde liegen. Diese nach der oben dargestellten Kategorisierung manipulierten bzw. nicht-authentischen Bewertungen werden in den nach folgenden Abschnitten F.3. und F.4. diskutiert. Nur wenige Entscheidungen stellen für die Beurteilung nach dem Lauterkeitsrecht allein darauf ab, dass die in Frage stehenden Bewertungen nicht als Produkttest gekennzeichnet sind:

So kam das LG München I in einem Urteil gegen einen Bewertungsvermittler zu dem Ergebnis, dass bereits dessen Verhalten als solches unlauter sei, da es ersichtlich eine irreführende geschäftliche Handlung im Sinne des § 5 Abs. 1 UWG darstelle.¹⁰⁷ Die Bewertungen seien als unwahre Angaben geeignet, Verbraucher, die die Plattform der Klägerin besuchten, zu geschäftlichen Handlungen zu veranlassen, welche sie bei einer wahren Bewertung nicht vorgenommen hätten.

¹⁰⁵ So beispielsweise LG Berlin, Beschl. v. 20.11.2018, Az. 15 O 203/18 (nicht veröffentlicht).

¹⁰⁶ So führt das OLG Frankfurt/M., Beschl. v. 22.02.2019, Az. 6 W 9/19 – *gekaufte Kundenbewertungen* = juris Rn. 23, aus, dass die Rezensenten bei Produkttests, wenn sie nicht schon selbst unternehmerisch handelten, jedenfalls zugunsten des Bewertungsvermittlers handeln und daher nicht lediglich als Tatmittler, sondern selbst als Täter anzusehen seien.

¹⁰⁷ LG München I, Urt. v. 28.03.2019, Az. 3 HK O 14073/18 (nicht veröffentlicht).

Das Oberlandesgericht Frankfurt/M. erkannte in einem Urteil gegen einen Bewertungsvermittler auf einen Verstoß gegen § 5a Abs. 6 UWG: Das Veröffentlichenlassen der Kundenbewertungen verfolge einen kommerziellen Zweck im Sinne des § 5a Abs. 6 UWG, der nicht kenntlich gemacht werde.¹⁰⁸ Der beklagte Bewertungsvermittler sei als Mittäter dafür verantwortlich, dass die von ihm beauftragten Produkttester ihre Bewertungen, ohne auf den kommerziellen Charakter hinzuweisen, auf der Seite des Portals einstellten. Ein Zusammenwirken zwischen dem Bewertungsvermittler und den von ihm beauftragten Produkttestern liege hier auf der Hand. Das gesamte Geschäftsmodell des Bewertungsvermittlers basiere schließlich darauf, dass die Produkttester-Bewertungen auf dem betroffenen Portal eingestellt würden, was sich schon aus dem Internet-Auftritt des Bewertungsvermittlers ergebe.¹⁰⁹ Der durchschnittlich informierte, situationsadäquat aufmerksame und verständige Durchschnittsverbraucher, so das Gericht, gehe bei Produktbewertungen grundsätzlich davon aus, dass diese ohne Gegenleistung erstellt werden, dass die Nutzer ihre Produkte aufgrund eines eigenen Kaufentschlusses erworben haben und ihre Bewertung unbeeinflusst von Dritten mitteilen. Der kommerzielle Zweck ergebe sich auch nicht unmittelbar aus den Umständen.¹¹⁰

Will man die verdeckten Produkttest-Bewertungen substantiell reduzieren, ohne den grundsätzlich gegebenen Wert von Produkttest-Bewertungen aufzugeben, muss man andere Wege gehen als bisher. Während die bisherige Rechtsdurchsetzung das Problem nicht nachhaltig lösen konnte, käme man nach vorläufiger Einschätzung des Bundeskartellamts durch eine stärkere Inpflichtnahme der Portale bei der Identifizierung von verdeckten Produkttest-Bewertungen oder durch die Möglichkeit der Kennzeichnung Produkttest-Bewertungen als solche zu sachgerechten Ergebnissen. Die Ermittlungen haben zahlreiche Indizien dafür ergeben, dass Portale bei der Bekämpfung von verdeckten Produkttest-Bewertungen nach wie vor auf Schwierigkeiten bei der Durchsetzung ihrer Maßnahmen stoßen. Die Gerichtsverfahren, die mehrere Portalbetreiber angestrengt haben, konnten die Menge verdeckter Produkttest-Bewertungen offenbar nicht alleine spürbar eindämmen. Unterlassungserklärungen wurden nicht in allen Fällen befolgt. Amazon geht zwar gegen solche Wiederholungstäter im Wege von Ordnungsgeldern vor. Verschiedene Marktteilnehmer haben dennoch die Einschätzung vertreten, dass die Zahl verdeckter Produkttests eher zu- als abnimmt.

Bei den Ermittlungen hat sich außerdem der Eindruck ergeben, dass im Verlauf des letzten Jahres Bewertungsvermittler weniger offen im Internet ihre Tätigkeit angeboten haben, sondern in (soziale) Netzwerke abgewandert sind bzw. ihren Sitz ins Ausland verlegt haben. Manche Bewertungsvermittler versuchen zwar auch, den Vorgaben der Rechtsprechung und der Plattformen nachzukommen und beschränken sich darauf, Produkttester zu vermitteln, diese aber nicht zum Einstellen ihrer Rezension auf Amazon zu verpflichten. Die

¹⁰⁸ OLG Frankfurt/M., Beschl. v. 22.02.2019, Az. 6 W 9/19, – *gekaufte Kundenbewertungen*, = juris Rn. 25.

¹⁰⁹ OLG Frankfurt/M., Beschl. v. 22.02.2019, Az. 6 W 9/19 – *gekaufte Kundenbewertungen*, = juris Rn. 18, 20. Das Gericht weist in seiner Entscheidung explizit darauf hin, dass die Anwendung des § 5a Abs. 6 UWG nicht durch die vorrangige Regelung des § 3 Abs. 3 i. V. m. Nr. 11 des Anhangs zum UWG verdrängt werde, da dieser vorliegend nicht einschlägig sei. Ein redaktioneller Inhalt im Sinne der Nr. 11 des Anhangs zum UWG liege nicht vor, da dieser ein Medium zur Verbreitung von Informationen voraussetze. Dazu könnten zwar auch elektronische Medien gehören, nicht aber sogenannte Bewertungsportale oder Foren im Internet, die lediglich ein Forum für die Bewertung von Waren oder Dienstleistungen durch Verbraucher bereitstellen (Rn. 14).

¹¹⁰ OLG Frankfurt/M., Beschl. v. 22.02.2019, Az. 6 W 9/19 – *gekaufte Kundenbewertungen*, = juris Rn. 25, 27 28.

Zahl der über diese Unternehmen vermittelten und auf großen Portalen verdeckt dargestellten Produkttest-Bewertungen hat dementsprechend abgenommen. Dies scheint aber durch die Zunahme der Vermittlungen über Facebook, WhatsApp und Telegram mehr als kompensiert zu werden.

Die bestehenden Urteile in dem Bereich zeigen aber gleichzeitig, dass der geltende Rechtsrahmen grundsätzlich ausreicht, um verdeckte Produkttests als verbraucherschädigende Irreführung zu ahnden. Die Einführung eines Per-se-Verbots in diesem Bereich, wie es mit Umsetzung der neuen Nr. 23c von Anhang I der UGP-Richtlinie zu erwarten ist (s. Kapitel E.4.2.), ist insofern eine hilfreiche Klarstellung, dürfte die bestehenden Probleme aber nicht lösen. Die Vorschrift in Nr. 23c UGP-Richtlinie definiert als unlautere Geschäftspraktik u. a. die Abgabe gefälschter Bewertungen bzw. die Erteilung des Auftrags an andere juristische oder natürliche Personen, gefälschte Bewertungen abzugeben. Diese Sachverhalte waren in der Vergangenheit schon unter § 5 Abs. 1 UWG bzw. § 5a Abs. 6 UWG erfasst worden, wie die oben zitierten Urteile zeigen. Der Grund für den Fortbestand des Problems verdeckter Produkttest-Bewertungen dürfte daher eher im Bereich der *Rechtsdurchsetzung* liegen und zwar nicht nur gegenüber unseriösen Bewertungsvermittlern, sondern auch gegenüber Portalen.

Im Folgenden sollen zwei Möglichkeiten aufgezeigt werden, die Flut verdeckter Produkttests einzudämmen: Zum einen könnten Portale mehr als bisher in die Pflicht genommen werden, gegen derartige Täuschungen vorzugehen. Zum anderen könnten Portale die Möglichkeit schaffen, offen als solche gekennzeichnete Produkttests-Bewertungen auf ihrer Seite anzuzeigen, um damit das Angebot unseriöser Bewertungsvermittler unattraktiver zu machen.

F.2.2. Lösungsansatz 1: Stärkere Inpflichtnahme der Portale

Portale sollten stärker in die Pflicht genommen werden, verdeckte Produkttestbewertungen auf ihren Seiten zu identifizieren und zu entfernen. Dies gibt der derzeitige gesetzliche Rahmen auch her.

Portale könnten mehr als bisher gegen verdeckte Produkttest-Bewertungen unternehmen, haben dazu derzeit aber nur bedingt einen Anreiz. Solange ein Missbrauch bzw. eine Gefährdung ihres Bewertungssystems nicht offensichtlich ist, haben sie von sich aus kein Interesse, aufwändige und möglicherweise kostenintensive Maßnahmen zu ergreifen, um verdeckte Produkttest-Bewertungen zu verhindern. Denn im Erfolgsfall führen diese Maßnahmen dazu, dass weniger Bewertungen auf dem jeweiligen Portal zu sehen sind. Dies ist eben nicht im Interesse eines Portals, denn viele Bewertungen machen es für Verbraucher attraktiv, führen so zu mehr Besuchen auf der Seite und damit höheren Umsätzen (s. Kapitel E.4.1.). Ein Portal dürfte daher erst dann Interesse entwickeln, Produkttest-Bewertungen effektiv zu bekämpfen, wenn seine Nutzer beginnen, an der Glaubwürdigkeit der Bewertungen auf dem Portal insgesamt zu zweifeln, wenn das Portal selbst in den Markt für Produkttests eintreten möchte oder wenn das Portal befürchten muss, für die durch die Irreführung entstehenden Schäden selbst in die Haftung zu geraten.

Die Zweifel der Nutzer können das Problem nicht lösen: Zwar sind z. B. bei Amazon in der Vergangenheit für viele Produkte Bewertungen angezeigt worden, die offensichtlich keine klassischen Bewertungen zu sein

schiene, und es werden Tipps im Internet veröffentlicht, wie man diese erkennen kann.¹¹¹ Diese Tipps legen nahe, dass man als Nutzer selbst sog. Fälschungen identifizieren und dann unberücksichtigt lassen kann. Tatsächlich ist aber davon auszugehen, dass Nutzer auch mit diesen Methoden gut gemachte verdeckte Produkttest-Bewertungen nicht erkennen, sondern sich möglicherweise sogar in einer falschen Sicherheit wiegen, wenn sie gewisse Bewertungen als „Fakes“ identifiziert haben. Sie vertrauen dann umso mehr auf die übrigen Bewertungen, ohne zu erkennen, dass es sich dabei ebenfalls um verdeckte Produkttests handeln könnte, die nur besser getarnt sind. Selbst bei einem Portal wie Amazon haben die öffentlichen Diskussionen um Fälschungen nicht dazu geführt, dass Verbraucher sich von dem Portal abwenden, sondern nur, dass sie misstrauisch gegenüber offensichtlichen Fälschungen geworden sind.

Nach den Ermittlungen ist Amazon ein Portal, das sehr entschlossen nicht nur gegen verdeckte Produkttests auf seiner Plattform vorgeht, sondern gegen Produkttests überhaupt. Das zeigt sich z. B. an der vergleichsweise großen Zahl von Gerichtsverfahren gegen Bewertungsvermittler, die Amazon angestrengt hat. Ein Grund dafür könnte sein, dass Amazon mit Vine ein eigenes Produkttesterprogramm betreibt und daher einen Anreiz hat, mit entsprechenden Gerichtsverfahren von Amazon unabhängige Bewertungsvermittler als alternative Anbieter von Produkttests vom Markt zu verdrängen.¹¹² Zumindest bisher haben allerdings Zugangsbeschränkungen zum Vine-Programm sowie die damit verbundenen hohen Kosten dazu geführt, dass weiterhin viele Händler auf verdeckte Produkttests setzen. Inwiefern Amazon alle Möglichkeiten ausschöpft, diese zu unterbinden, ist unklar. Zwar hat Amazon von allen befragten Portalen die meisten Verfahren gegen Bewertungsvermittler angestrengt. Verschiedene Marktteilnehmer haben aber ausgeführt, dass auch Amazon mehr unternehmen könnte, um gegen Produkttests gerade bei einer Vermittlung über (soziale) Medien vorzugehen.

Was die durch die Irreführung entstehenden Schäden angeht, müssen Portale derzeit praktisch nicht befürchten, für diese haftbar gemacht zu werden. Dabei ist im Einklang mit der Leitentscheidung „*Hotelbewertungsportal*“ des Bundesgerichtshofs¹¹³ in einigen der untersuchten Konstellationen davon auszugehen, dass die betreffenden Portale ihren diesbezüglichen Verkehrspflichten nicht ausreichend nachkommen. Der Bundesgerichtshof führt dort folgendes aus: Der Haftung wegen Verletzung wettbewerbsrechtlicher Verkehrspflichten liege der Gedanke zugrunde, dass derjenige, der in seinem Verantwortungsbereich eine Gefahrenquelle schaffe oder andauern lasse, die ihm zumutbaren Maßnahmen und Vorkehrungen treffen müsse, die zur Abwendung der Dritten daraus drohenden Gefahren notwendig seien.¹¹⁴ Eine allgemeine Überwachungsspflicht von Diensteanbietern, zu denen auch Internet-Bewertungsportale gehörten, sei aufgrund der auch als Plattformprivileg bezeichneten Vorschriften des TMG allerdings ausgeschlossen. Eine solche allgemeine

¹¹¹ So z. B. <https://www.mdr.de/nachrichten/ratgeber/digital-technik/amazon-falsche-kunden-bewertungen-fake-be-trug-test-100.html>, <https://www.pc-magazin.de/ratgeber/amazon-bewertungen-pruefen-fake-rezensionen-erkennen-tipps-online-shopping-3200496.html> (beide abgerufen am 01.10.2020).

¹¹² S. dazu auch die kritischen Äußerungen des OLG Frankfurt/M., Beschluss vom 22.02.2019, Az. 6 W 9/19, – *gekaufte Kundenbewertungen* = juris Rn. 30.

¹¹³ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, – *Hotelbewertungsportal*.

¹¹⁴ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 42 – *Hotelbewertungsportal*.

Überwachungspflicht würde verlangen, dass Portale grundsätzlich jeden Beitrag auf eine mögliche Rechtsverletzung hin prüfen müssten. Deutlich weniger weitgehend und durchaus mit den auf europäischen Vorgaben beruhenden Vorschriften des TMG vereinbar sei dagegen eine spezifische Überwachungspflicht. Diese richte sich danach, ob dem Bewertungsportal nach den Umständen eine Prüfung zuzumuten ist. Dabei sei zu berücksichtigen, wie aufwändig eine derartige Prüfung wäre und ob der Betreiber bei seiner Tätigkeit Rechtsverletzungen in erheblichem Umfang Vorschub leistet oder sie durch eigene Maßnahmen fördert.¹¹⁵

Gemessen daran hält das Bundeskartellamt verstärkte Maßnahmen der Portale gegen verdeckte Produkttest-Bewertungen für zumutbar: In Anbetracht ihres massenhaften Auftretens handelt es sich bei den verdeckten Produkttest-Bewertungen um vorhersehbare Rechtsverstöße. Schon jetzt übernehmen Portale für die Entfernung von vergleichbar vorhersehbaren Rechtsverstößen wie z. B. für Diskriminierungen, Hate Speech oder Datenrechtsverstöße in Rezensionstexten ganz selbstverständlich die Verantwortung, obgleich sie ganz überwiegend nicht zum Kreis der Normadressaten des NetzDG gehören. Für eine entsprechende Filterung könnten Portale z. B. Metadaten der Bewertungsabgabe analysieren oder automatisiert Verbindungen zwischen Bewertungen, Angeboten von Bewertungsvermittlern und Rezensenten herstellen. Einzelne Portale wenden auch diese Methoden bereits an. (s. Kapitel D.2.1. und D.2.2.). Entsprechend ist in der Fallkategorie der verdeckten Produkttestbewertungen nach vorläufiger Einschätzung von einer spezifischen Überwachungspflicht für Portale auszugehen.

Außerdem gibt es Fälle, in denen Portale darüber hinaus durch die Ausgestaltung ihres Portals der Entstehung verdeckter Produkttest-Bewertungen Vorschub leisten. Das geschieht z. B. dadurch, dass sie keine Möglichkeit schaffen, Produkttest-Bewertungen zur Kennzeichnung offen zu legen, oder durch einen hohen Einfluss von Bewertungen auf die Auffindbarkeit und das Ranking von Produkten auf ihrer Seite. Ein derartiges Verhalten erhöht nach höchstrichterlicher Rechtsprechung die Verantwortlichkeit eines Portals: Je mehr es durch sein Verhalten zum Auftreten verdeckter Produkttest-Bewertungen beiträgt, desto höher sind die Anforderungen an die spezifischen Überwachungspflichten des Portalbetreibers.¹¹⁶ Diese spezifischen Überwachungspflichten können jedenfalls dann nicht als erfüllt angesehen werden, wenn es schon Dritten jederzeit möglich ist, höchstwahrscheinlich auf verdeckte Produkttests zurückzuführende Bewertungen auf den Seiten eines Portals zu finden. Dies sind z. B. Bewertungen für mehrere Produkte, die jeweils in (sozialen) Netzwerken zum Test angeboten wurden und die jeweils von der gleichen Person bewertet wurden.

Die Befolgung solcher spezifischen Überwachungspflichten birgt nicht die Gefahr, dass das Portal sich die Bewertungen durch die entsprechenden Überprüfungen zu Eigen macht. Schon jetzt verwenden Portale Wortfilter, die Schimpfwörter filtern, oder beseitigen Datenschutzverstöße. Nicht anders wäre auch die Überprüfung auf verdeckte Produkttests auf die Einhaltung der Nutzungsbedingungen und eigener Rechtspflichten ausgerichtet, nicht auf eine inhaltliche Kontrolle der Bewertungen auf Richtigkeit.¹¹⁷

¹¹⁵ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 31, 36 – *Hotelbewertungsportal*.

¹¹⁶ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 36 – *Hotelbewertungsportal*.

¹¹⁷ Vgl. BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 28 – *Hotelbewertungsportal*; Spindler in: Spindler/Schmitz, Kommentar zum TMG, 2. Auflage 2018, § 10, Rn. 93.

Aus den gleichen Gründen geht die Argumentation fehl, dass eine Verantwortlichkeit der Portale schon aufgrund des europarechtlich verankerten Plattformprivilegs nicht möglich sei. Schon jetzt verhindert das Plattformprivileg nicht den überwiegend freiwilligen Einsatz von Filtern gegen Hate Speech, Diskriminierungen u. a. Die Rechtsprechung sollte zudem bei der Auslegung des Plattformprivilegs berücksichtigen, dass das Plattformprivileg nur für Fälle gedacht ist, in denen die Tätigkeit des Anbieters „rein technischer, automatischer und passiver Art“ (Erwägungsgrund 42 ECLR-Richtlinie) ist. Das ist beim Betrieb eines Bewertungssystems in der Regel nicht der Fall. Im Unterschied z. B. zu einem Cloud-Anbieter, für dessen Geschäft völlig irrelevant ist, in welcher Form welche Inhalte dort abgelegt werden, sind für ein Bewertungsportal die Bewertungen Inhalte, die den eigenen Geschäftserfolg maßgeblich beeinflussen. So kann es für ein Portal z. B. wichtig sein, Restaurants in einem bestimmten Gebiet flächendeckend mit Bewertungen abbilden zu können, oder gute und schlechte Bewertungen anzeigen zu können, um einen möglichst hohen Informationsgehalt für die Nutzer bieten zu können. Die abgegebenen Bewertungen werden auf dem Portal daher auch nicht einfach abgelegt, sondern von dem Portal selbst möglichst nutzerfreundlich dargestellt. Oft wird eine von anderen Nutzern als besonders hilfreich markierte Bewertung besonders prominent platziert, die Reihenfolge der Bewertungen ist ebenso wenig zufällig wie ihre Platzierung auf der Seite und die zusammenfassende Darstellung aller Bewertungen. Die Verfasser von Bewertungen erhalten teilweise von den Portalen (virtuelle) Prämien nicht nur für viele, sondern auch für inhaltlich besonders hilfreiche Bewertungen (s. Kapitel D.1.3.). All dies zeigt, dass der Inhalt der Bewertungen inhärenter Teil des Geschäftsmodells von Portalen ist und dass dies bei der Auslegung des Plattformprivilegs nicht unberücksichtigt bleiben kann.

Wenn Hinweise darauf bestehen, dass Portale ihren diesbezüglichen Überwachungspflichten nicht nachkommen, sollten in Ermangelung behördlicher Eingriffsbefugnisse in Bezug auf das UWG Verbraucherschutzorganisationen versuchen, dies auf dem Rechtswege einzuklagen. Denn Verbraucher haben ein eindeutiges und originäres Interesse daran, gerade auch nicht offensichtliche, schwer oder gar nicht erkennbare verdeckte Produkttest-Bewertungen als Irreführungstatbestände rechtlich zu ahnden.

In der Konsultation zur vorliegenden Sektoruntersuchung wurde eine stärkere Inpflichtnahme der Portale naturgemäß unterschiedlich bewertet: Der vzbv begrüßte den Vorschlag, Portale an verschiedenen Stellen stärker in die Verantwortung zu nehmen, ausdrücklich. Dies sei für die hier betrachteten Portale, auf denen Produkttestbewertungen eine wesentliche Rolle spielen (können), gerechtfertigt und zumutbar.¹¹⁸ Ein Portal wies darauf hin, dass die Identifizierung von Fake-Bewertungen äußerst komplex und daher für die Portale nicht einfach sei. Ein anderes Portal wies darauf hin, dass zumindest allgemeine Bewertungsplattformen nach geltender Rechtsprechung unter das Plattformprivileg fallen und daher eine stärkere Inpflichtnahme zumindest für diese Plattformen nach geltendem Recht ausgeschlossen sei.

¹¹⁸ Auf Interaktionsplattformen wie YouTube, Twitter oder Facebook, bei denen die Kuratierung nutzergenerierter medialer Inhalte sowie die Bereitstellung und Vermittlung derselben im Mittelpunkt stehe, sei eine stärkere Verantwortung von Portalen für Inhalte eher kritisch zu sehen. Siehe vzbv, Verantwortung von Online-Plattformen, 2020, Verbraucherpolitische Vorschläge zur Novellierung der haftungsrechtlichen Regelungen der E-Commerce-Richtlinie im Rahmen des Digital Services Act (DSA).

F.2.3. Lösungsansatz 2: Kennzeichnung von Produkttest-Bewertungen

Ein zweiter und möglicherweise erfolgsversprechenderer Weg, die Menge verdeckter Produkttestbewertungen zu verringern, besteht darin, dass Portale Möglichkeiten schaffen, auf Produkttest basierende Bewertungen zu kennzeichnen und damit aus der Illegalität zu holen. Bisher gehen Versuche, verdeckte Produkttest-Bewertungen zu reduzieren, fast immer in die Richtung, dass entsprechende Bewertungen insgesamt verhindert werden sollen. Eine sinnvolle Alternative oder zumindest Ergänzung zu diesen Maßnahmen besteht darin, klar als Produkttest gekennzeichnete Bewertungen auf einem Portal zuzulassen. Derart offene Produkttest-Bewertungen stellen keinen Lauterkeitsrechtsverstoß dar, weil der Nutzer des Portals damit über das Zustandekommen der Bewertung informiert ist.

Bewertungen von Produkttestern sollten zusammen mit denen anderer Verbraucher angezeigt werden. Dabei muss für jede einzelne Bewertung klar erkennbar sein, ob diese aufgrund eines Produkttests abgegeben wurde oder nicht. Auch bei der zusammenfassenden Darstellung sollten Bewertungen von Produkttestern berücksichtigt werden können. Dabei muss ebenfalls transparent sein, ob und in welcher Form Produkttest-Bewertungen mit einbezogen werde. So könnte eine Zusammenfassung unter Ausschluss von Produkttest-Bewertungen voreingestellt sein, dem Verbraucher aber die Möglichkeit gegeben werden, sich alternativ einen Durchschnitt über alle Bewertungen oder nur über alle Produkttest-Bewertungen anzeigen zu lassen. Zusätzliche Transparenz kann durch erläuternde Texte, die interessierten Verbrauchern bei Bedarf zugänglich sind, geschaffen werden.


Ein nicht gekennzeichneter Produkttest stellt einen Verstoß gegen § 5a Abs. 6 UWG dar, denn es handelt sich dabei um getarnte Werbung: Die auf Absatzförderung gerichtete Verhaltensweise des Anbieters ist für den Verbraucher nicht klar, verständlich und in eindeutiger Weise erkennbar.¹¹⁹ Eine von einem Produkttester verfasste Bewertung müsste daher ebenfalls in geeigneter Form als solche gekennzeichnet sein. Eine Kennzeichnung, die den Charakter des Produkttests erkennbar macht, wäre dabei der alternativ denkbaren, allgemeineren Kennzeichnung als „Werbung“ vorzuziehen: Der Begriff der Werbung umfasst allgemein alle Formen von Informationen, denen eine auf Absatzförderung gerichtete Verhaltensweise des Anbieters zugrunde liegt. Ein Produkttest stellt dagegen eine sehr spezielle Form der Absatzförderung bzw. Werbung dar, bei der der Initiator auf jegliche inhaltliche Kontrolle verzichtet. Der Anbieter hat keine Möglichkeit, einmal in Auftrag gegebene Bewertungen eines Produkttesters zu beeinflussen, und er hat auch keine Möglichkeit, ihre Darstellung auf dem Bewertungsportal zu verhindern. Der Anbieter setzt sich mit der Beauftragung von Produkttests daher einem Risiko aus, das er sonst in dieser Form bei Werbung nicht trägt, nämlich dass sein Produkt in der Öffentlichkeit unumkehrbar negativ dargestellt wird. Für den Verbraucher enthält die Kennzeichnung als „Produkttest“ daher gegenüber einer allgemeinen Kennzeichnung als „Werbung“ eine zusätzliche Information, die zusätzliche Transparenz und ggfs. ein größeres Vertrauen in die enthaltenen Aussagen schafft.

¹¹⁹ Büscher in: Büscher, Kommentar zum UWG, § 5a, Rn. 189 f.

Die Bewertungsvermittler sind dabei verantwortlich dafür, sicher zu stellen, dass Anbieter als Auftraggeber von Produkttests keinen Einfluss auf den Inhalt der Bewertung nehmen können, sondern höchstens formale Vorgaben wie zur Länge einer Bewertung möglich sind. Bewertungsvermittler können dies z. B. durch die Art der Vermittlung sicherstellen, indem sie den Versand des Testprodukts an den Produkttester selbst übernehmen. Die Portale, die die Möglichkeit anbieten, Produkttests auf ihrer Seite einzustellen, sind verantwortlich dafür, dass Bewertungen von Produkttestern, die gleichwertig neben den klassischen Bewertungen dargestellt werden, auch nach den gleichen Kriterien gefiltert und ebenso wie andere Bewertungen im Nachhinein nur im Falle von Verstößen gegen Bewertungsrichtlinien modifiziert oder gelöscht werden können.

Der vzbv hat im Rahmen seiner Stellungnahme zum Konsultationspapier darauf hingewiesen, dass damit zu rechnen sei, dass auch ohne eine explizite Einflussnahme des Anbieters Produkttester dazu neigen würden, Items tendenziell positiver zu bewerten als ohne materielle Gegenleistung, um vermeintlich unausgesprochene Erwartungen zu erfüllen und bei künftigen Auftragsvergaben nicht benachteiligt zu werden. In der Tat wäre es nicht überraschend, wenn die Sterne-Bewertungen von Produkttestern im Durchschnitt etwas besser ausfallen als die von anderen Bewertungsverfassern, auch wenn ein Bewertungsvermittler explizit darauf hingewiesen hat, dass dies bei den von ihm beauftragten Produkttestern nicht der Fall sei. Durch die transparente Unterscheidung der zwei Herkunftsarten von Bewertungen in der Einzeldarstellung und in der zusammenfassenden Darstellung kann der Verbraucher sich aber selbst ein Bild machen und z. B. bei stark abweichenden Durchschnittswerten seine eigenen Schlüsse daraus ziehen. Es ist im Übrigen auch denkbar, dass die Produkttester sich mit dem Inhalt ihrer Bewertungen mehr Mühe als sonstige Bewertungsverfasser geben, um den Erwartungen ihrer Auftraggeber gerecht zu werden. Das nutzt wiederum den Verbrauchern, für die diese sorgfältig verfassten Bewertungen entsprechend hilfreicher sein können als klassische Bewertungen.


Die Möglichkeit, Bewertungen von Produkttestern zu kennzeichnen, bieten Portale bisher nur vereinzelt an. So hat z. B. Amazon hat mit dem Programm Vine eine Möglichkeit der Kennzeichnung geschaffen, allerdings nur für selbst vermittelte Produkttests. Der Nutzer wird darüber durch den in grüner Farbe markierten Hinweis „Vine Kundenrezension eines kostenfreien Produkts“ informiert. Weitere Informationen erhält der Nutzer, wenn er den blau markierten Text „Was ist das?“ anklickt (s. Abb. unten).


 Feedback für dich 🍌 **TOP 500 REZENSENT**

★★★★★ **Traumhaft schöne Einladungskarten**
 Rezension aus Deutschland vom 21. August 2020
Vine Kundenrezension eines kostenfreien Produkts (Was ist das?)

Diese 10 Einladungskarten mit ihren rosafarbenen Umschlägen sind ja wohl wirklich ein Traum für jeden Einhornfan. Mit goldener Umrandung und Sternen und einem wunderschönen Einhorn in der Mitte - einfach nur toll. Dazu gefällt mir die gewählte Schriftart ausgesprochen gut. Sie passt einfach perfekt zum restlichen Design. Sowohl die Druck- als auch die Materialqualität ist angemessen und fehlerfrei.


Durch den vorgedruckten Lückentext sind die Einladungen schnell erstellt und jedem Gast ist klar: Was, wo und wann stattfinden wird und das er doch bitte Rückmeldung geben soll. Auch für sonstige Wünsche ist noch ein bisschen Platz ganz am Ende.

 **Fazit:**
 Wunderschöne Einladungskarten und durch das Design auch ganz klar noch für etwas ältere Einhornfans geeignet.



Eine Person fand diese Informationen hilfreich

| |

 **Sciurus**

★★★★☆ **Teilweise zu kurze Lücken / Tinte verschmiert stark / nassklebende, durchscheinende Umschläge**
 Rezension aus Deutschland vom 30. August 2020
Vine Kundenrezension eines kostenfreien Produkts (Was ist das?)

Die Karten sind sehr stabil und haben auf der Bildseite eine glatte und auf der Schreibseite eine etwas rauere Oberfläche, was mir sehr gut gefällt. Schön finde ich auch die abgerundeten Ecken, was diese Karten hochwertiger wirken lässt.

Nachdem der erste Eindruck rundweg positiv war, folgte die Ernüchterung beim Ausfüllen der Karten. Die Tinte von Finelinern und Füllern (zum Teil sogar von Kugelschreibern) trocknet sehr schlecht, sodass noch bis zu einer halben Minute oder länger die Gefahr besteht, dass das Geschriebene verwischt. Gerade Kinder (und Linkshänder) müssen hier also besonders vorsichtig sein.

Mal davon abgesehen, dass es mich persönlich stört, dass auf der Karte das Wörtchen „bis“ groß geschrieben ist, obwohl es mitten im Satz steht, sind manche Lücken schon recht kurz gehalten. Eine vollständige E-Mail-Adresse

Abbildung 16: Vine-Kundenrezensionen bei Amazon¹²⁰

Externe haben auf Amazon keine Möglichkeit, Bewertungen von Produkttestern einzustellen und entsprechend zu kennzeichnen, da diese vom System gelöscht werden. Die Masse an Angeboten für Amazon-Bewertungen im Internet und in (sozialen) Netzwerken zeigt, dass der Bedarf von Produkttests bei Amazon besonders groß zu sein scheint. Eine Öffnung des Portals für Produkttests, die von Externen vermittelt werden, würde Anbietern auf Amazon mehr Möglichkeiten eröffnen, Produkttests in Auftrag zu geben und Verbraucher damit über das eigene Produkt zu informieren.

Es ist davon auszugehen, dass die Möglichkeit, Produkttests auf legalem Weg und zu wettbewerblichen Konditionen auf Plattformen zu posten, von Anbietern, die darauf angewiesen sind, rege genutzt würde. Das

¹²⁰ https://www.amazon.de/artpin%2CAE-EINLADUNGSKARTEN-Einschulung-Kindergeburtstag-Umschl%2CA4gen/dp/B08DG7X53N?ref=pf_vv_at_pdctrvw_dphttps://intranet/DE/Home/home_NODE.html (abgerufen am 01.10.2020).

haben insbesondere die Aussagen und Stellungnahmen der Bewertungsvermittler im Rahmen der Sektoruntersuchung ergeben. Das Bundeskartellamt teilt außerdem die Einschätzung dieser Marktteilnehmer, dass durch eine solche Möglichkeit die Zahl verdeckter Produkttest-Bewertungen von selbst deutlich reduziert würde. Viele der Anbieter, die jetzt verdeckte Produkttest-Bewertungen in Auftrag geben, tun dies, weil sie es für den Erfolg ihres Geschäfts für unabdingbar halten, und riskieren dafür ein rechtlich fragwürdiges Verhalten. Gäbe es für diese Anbieter eine rechtlich sichere Alternative, würden sie diese nutzen, auch wenn deren Wirksamkeit möglicherweise etwas geringer ist. Nicht vergessen werden darf dabei, dass für die Mehrzahl aller Produkte Produkttests nicht notwendig sind, weil diese ohnehin schon viele Bewertungen aufweisen. Aufgrund der Kosten und des Aufwands für Produkttests würden diese vermutlich vor allem für besondere Produkte wie Neuheiten eingesetzt, bei denen noch keine oder wenige Bewertungen vorliegen. Es ist also nicht zu erwarten, dass klassische Nutzerbewertungen in größerem Umfang von Produkttests verdrängt werden.

F.3. Manipulierte Bewertungen

Der vorige Abschnitt F.2. beschäftigt sich mit Produkttest-Bewertungen, deren Inhalt unbeeinflusst vom Anbieter bzw. Auftraggeber ist. Viele Bewertungsvermittler haben dem Bundeskartellamt gegenüber betont, dass sie darauf achten, Produkttester inhaltlich in keiner Weise zu beeinflussen, abgesehen möglicherweise von formalen Vorgaben wie z. B. einer Mindestlänge für Rezensionen. Gegenüber dem Auftraggeber könne daher immer nur eine bestimmte Anzahl an Bewertungen insgesamt garantiert werden, nicht aber eine Anzahl an 5-Sterne-Bewertungen. Es gibt jedoch auch andere Geschäftsmodelle. Bei einigen Bewertungsvermittlern wird die Vermittlung von Produkttests mit inhaltlichen Vorgaben verbunden wie z. B. einer Bewertung von mindestens vier oder fünf Sternen. Bei diesen Bewertungen erfolgt die Nutzung des zu bewertenden Items wie bei den im vorherigen Abschnitt diskutierten Produkttests ebenfalls nicht aufgrund einer eigenen Entscheidung, sondern beruht auf einem Anreiz. Entscheidend ist jedoch, dass in diesen Fällen darüber hinaus auch der Inhalt der Bewertung nicht frei und unbeeinflusst zustande kommt, sondern durch Vorgaben des Auftraggebers manipuliert wird (s. Abb. 16). Nicht anders als bei den Produkttest-Bewertungen dürften die Auftraggeber für manipulierte Bewertungen im Regelfall Anbieter auf Plattformen sein, die damit ihre eigenen Angebote gegenüber denen anderer Anbieter vorteilhafter positionieren möchten.

F.3.1. Ökonomische und rechtliche Bewertung

Manipulierte Bewertungen sind offensichtlich ökonomisch schädlich: Sie führen dazu, dass Verbraucher schlechtere und/oder teurere Produkte kaufen oder die falschen Dienstleister auswählen, weil diese aufgrund der Manipulation der Bewertungen besser als ihre Substitute zu sein scheinen. Manipulierte Bewertungen verzerren damit zugleich den Qualitäts- und Preiswettbewerb der Anbieter untereinander.

Dass manipulierte Bewertungen einen Verbraucherrechtsverstoß darstellen, ist ebenfalls unstrittig: Bereits nicht gekennzeichnete Produkttests stellen einen Verstoß jedenfalls gegen § 5 Abs. 1 Nr. 1 bzw. § 5a Abs. 6 UWG dar (s. Kapitel F.2.1.). Dies gilt erst recht für manipulierte Bewertungen, die gewissermaßen doppelt

gegen das Verbot der Irreführung verstoßen: Zunächst wird der kommerzielle Zweck der Bewertung nicht offengelegt, und darüber hinaus wird die Bewertung inhaltlich beeinflusst.

Eine Rechtsprechungspraxis ist bei manipulierten Bewertungen – vergleichbar mit den verdeckten Produkttest-Bewertungen – daraus entstanden, dass in den letzten Jahren einige Portale damit begonnen haben, die Vermittler von Bewertungen und in Einzelfällen auch die Anbieter, die diese Vermittlungen beauftragt hatten, zivilrechtlich zu verfolgen (s. Kapitel E.3.1.). Manche Gerichte haben sich bei der Begründung ihrer daraus entstandenen Entscheidungen allein auf die Existenz verdeckter Produkttests beschränkt, einige gehen aber auch explizit auf den teilweise darüber hinaus vorliegenden Aspekt der Manipulation ein. Gerichte halten eine Manipulation z. B. für nachgewiesen, wenn der Bewertungsvermittler mit einer „100 % Zufriedenheitsgarantie“ wirbt. Damit seien offensichtlich Bestbewertungen gemeint, und die könne man nur versprechen, wenn man Manipulationen einbezieht. Als klaren Hinweis auf Manipulation werten die Gerichte auch, wenn Bewertungsvermittler ihren Kunden das Angebot unterbreiten, vorgegebene Texte oder Bilder in die Rezensionen aufzunehmen.¹²¹ Das LG Hamburg führt in diesem Zusammenhang aus, dass eine Kundenrezension auch dann keineswegs unbeeinflusst erfolge, wenn nicht *expressis verbis* versprochen werde, dass sie durchwegs positiv ausfallen würde. Denn das Angebot des angeklagten Bewertungsvermittlers sehe vor, dass die Warenanbieter die Texte der Rezension mitliefern dürften und die Anbieter bei nicht hundertprozentiger Zufriedenheit ihr Geld zurückverlangen könnten. Es sei glaubhaft, dass eine solche „hundertprozentige Zufriedenheit“ insbesondere die Abgabe von fünf Sternen umfasse.¹²² Neben dem LG Hamburg stellt jedoch auch das LG München I¹²³ bei der Beurteilung eines quasi identischen Sachverhalts klar, dass ein Nicht-Kennlichmachen des kommerziellen Zweckes bereits ausreichend für einen Verbraucherrechtsverstoß sei.

Das Bundeskartellamt hat im Rahmen seiner Ermittlungen den Eindruck gewonnen, dass manipulierte Bewertungen in den letzten Monaten tendenziell zugenommen haben. Dies ist möglicherweise auf die verstärkten rechtlichen Bemühungen großer Portale zurückzuführen, gegen verdeckte Produkttests vorzugehen, auch wenn dies zunächst paradox erscheinen mag. Tatsächlich hat das Vorgehen der Portale gegen Bewertungsvermittler aber dazu geführt, dass die Vermittlung von Bewertungen zunehmend über inoffizielle, schwer zu verfolgende Wege wie in (sozialen) Netzwerken oder im Ausland stattfindet, wo Manipulationen aufgrund der geringeren Transparenz schwieriger nachzuweisen sind. Denn die Nachfrage der Anbieter nach derartigen Bewertungen ist unverändert hoch. Für seriöse Bewertungsvermittler ohne Manipulationsabsicht ist es über derartige Kanäle noch schwieriger, sich gegenüber Vermittlern zu profilieren, die mit Zufriedenheitsgarantien werben. Außerdem haben Vermittler angesichts der Tatsache, dass nicht manipulierte, verdeckte Produkttest-Bewertungen faktisch in gleicher Weise geahndet werden wie solche, die darüber hinaus auch noch manipuliert sind, erst recht keinen Anreiz, Manipulationen zu vermeiden.

¹²¹ So z. B. bei LG München I, Urt. v. 28.03.2019, Az. 3 HK O 14073/18 (nicht veröffentlicht).

¹²² LG Hamburg, Beschl. v. 23.08.2018, Az. 815 O 275/18 (nicht veröffentlicht).

¹²³ LG München I, Beschl. v. 04.08.2018, Az. 33 O 10931/18 (nicht veröffentlicht).

F.3.2. Lösungsansatz: Effektivere Rechtsdurchsetzung

Wie bereits bei den verdeckten Produkttests unter F.2.2. diskutiert, sollten neben der stärkeren Verfolgung unseriöser Bewertungsvermittler auch bei manipulierten Bewertungen Portale für deren Veröffentlichung stärker zur Verantwortung gezogen werden. Dies ist mit dem derzeitigen Rechtsrahmen vereinbar. Sie sollten außerdem die Möglichkeit schaffen, Produkttest-Bewertungen zu kennzeichnen, um diese aus der Illegalität zu holen und die Nachfrage nach manipulierten Bewertungen zu reduzieren. Dies würde möglicherweise dazu beitragen, dass Produkttest-Bewertungen weniger häufig manipuliert werden. Versuche, Tester zu beeinflussen, könnten dann möglicherweise einfacher identifiziert und dann verhindert werden.

Hauptsächlich müssen jedoch die geltenden rechtlichen Möglichkeiten ausgeschöpft werden, gegen verbraucherschädigende manipulierte Bewertungen vorzugehen. Für Bewertungsvermittler und ihre Auftraggeber dürfen das Abwandern in (soziale) Netzwerke und eine schwere Erreichbarkeit im Inland keine Möglichkeiten darstellen, sich geltendem Verbraucherrecht zu entziehen. Gefragt sind bei der Durchsetzung mangels entsprechender behördlicher Befugnisse in erster Linie die betroffenen Portale und private Verbraucherschutzorganisationen.

F.4. Nicht-authentische Bewertungen

Nicht-authentische Bewertungen unterscheiden sich von manipulierten Bewertungen dadurch, dass der Verfasser der Bewertung das zu bewertende Item überhaupt nicht selbst genutzt hat und ggfs. auch gar keine natürliche Person ist (s. Abb. 16). Eine objektive Bewertung ist damit von Vorneherein nicht möglich, denn ohne Nutzererfahrung hat ein Verfasser keine Möglichkeit, sich selbst eine Meinung darüber zu bilden, wie er ein Item bewertet. Eine nicht-authentische Bewertung ist immer auch manipuliert, weil davon auszugehen ist, dass der Verfasser mangels eigener Erfahrung eine Bewertung im Sinne seines Auftraggebers abgibt.

Auftraggeber für nicht-authentische Bewertungen kann ein Portal selber sein, das auf der eigenen Seite angebotene Items damit bewerben möchte. Im Regelfall dürften nicht-authentische Bewertungen aber von Anbietern auf Plattformen beauftragt worden sein, die damit ihre eigenen Angebote gegenüber denen anderer Anbieter auf der Plattform besser dastehen lassen möchten. In Einzelfällen können auch Anbieter und Bewerter identisch sein, wenn z. B. ein Anbieter selbst händisch Bewertungen für die eigenen Items abgibt oder durch einen Bot automatisiert Bewertungen für seine Angebote generieren lässt. Da derartige Selbstbewertungen aber, vor allem wenn sie massenhaft auftreten, von den meisten Portalen mit technischen Mitteln entdeckt und dann entfernt werden können, werden in der Regel Dritte mit dem Einstellen der entsprechenden Bewertungen beauftragt.

Bei der Bewertung von physischen Produkten spielen nicht-authentische Bewertungen eher keine Rolle: Nach den Ermittlungen sind die im Kapitel F.3. analysierten *manipulierten* Bewertungen ein Phänomen, das vorwiegend dort auftritt. Die „Bezahlung“ des Produkttesters erfolgt in der Regel dadurch, dass der Tester das Produkt umsonst oder stark ermäßig erhält. Hinweise auf *nicht-authentische* Bewertungen für Produkte (insbesondere durch Bots bzw. über „gekaperte“ Kunden-Accounts) hat das Bundeskartellamt lediglich in

Bezug auf technische Artikel von chinesischen Händlern erhalten. Grundsätzlich sind nicht-authentische Bewertungen eher ein Phänomen, das bei der Bewertung von Dienstleistungen auftritt. Obwohl diese derzeit weniger im Fokus der Öffentlichkeit stehen – in der Presse werden gefälschte Produktbewertungen insbesondere bei Amazon sehr viel stärker thematisiert –, spielen auch in diesem Bereich Fälschungen eine nicht zu vernachlässigende Rolle: Die in Kapitel E.2.1. vorgestellten ökonomischen Studien fanden Hinweise auf nicht-authentische Bewertungen im Gastronomie- und Hotelbereich auf zwei großen Bewertungsportalen. Manche Bewertungsvermittler bieten neben Bewertungen auf Reiseportalen auch die Vermittlung von Ärztebewertungen z. B. auf Jameda oder von Autohändlern z. B. auf Mobile.de an. Bei diesen Bewertungen ist davon auszugehen, dass derjenige, der die Bewertung abgibt, nicht selber den betreffenden Arzt, das Hotel oder den Autohändler aufsucht, sondern tatsächlich nur eine Bewertung dafür abgibt.

F.4.1. Ökonomische und rechtliche Bewertung

Nicht-authentische Bewertungen sind ökonomisch schädlich: Sie führen dazu, dass Verbraucher Waren oder Dienstleistungen auswählen, die teurer, schlechter oder weniger geeignet sind als ihre Bewertung suggeriert. Ehrliche Anbieter werden benachteiligt, der Qualitäts- und Preiswettbewerb der Anbieter untereinander wird verzerrt.

Nicht-authentische Bewertungen stellen in rechtlicher Hinsicht einen Verstoß gegen § 5 Abs. 1 Nr. 1 bzw. § 5a Abs. 6 UWG dar. Die Irreführung des Verbrauchers findet dabei verglichen mit den manipulierten Bewertungen und verdeckten Produkttests schon eine Stufe früher statt: Der Bewertung liegt überhaupt keine eigene Verbrauchererfahrung des Bewertungsverfassers zugrunde. Davon dürften Verbraucher beim Lesen von Bewertungen aber in jedem Fall ausgehen.

Das Ausmaß, in dem nicht-authentische Bewertungen auf Portalen zu finden sind, hängt nicht unerheblich von der Art des Bewertungssystems ab. Eine ganze Reihe von Portalen verwendet geschlossene Bewertungssysteme, bei denen nur Nutzer, die vom Portal eine entsprechende Aufforderung bekommen, eine, und zwar nur eine einzige, Bewertung abgeben können. Die Verwendung eines solchen Systems ist immer dann möglich, wenn das Portal die Vermittlung der zu bewertenden Items übernimmt und so Kenntnis darüber hat, welcher Nutzer eine Leistung tatsächlich in Anspruch genommen bzw. ein Produkt erworben hat. So können auf Booking und Expedia nur Kunden ein Hotel bewerten, die über das jeweilige Portal gebucht und dann auch tatsächlich die Übernachtung dort angetreten haben. Bei derartigen geschlossenen Bewertungssystemen kann es allerdings erwünscht sein, dass auch diejenigen Nutzer eine Bewertung abgeben können, bei denen am Ende keine provisionspflichtige Transaktion zustande kommt. Die Gründe für das Scheitern können sogar besonders informativ und relevant für andere Nutzer sein.¹²⁴ Grundsätzlich weisen Portale mit geschlossenen Bewertungssystemen tendenziell weniger und möglicherweise auch weniger repräsentative Bewertungen auf als solche mit offenen Bewertungssystemen. Damit stellt ein geschlossenes System zwar eine

¹²⁴ Bundeskartellamt, Sektoruntersuchung Vergleichsportale, S. 116 f., April 2019, https://www.bundeskartellamt.de/SharedDocs/Publikation/DE/Sektoruntersuchungen/Sektoruntersuchung_Vergleichsportale_Bericht.html (abgerufen am 01.10.2020).

wirksame Maßnahme gegen nicht-authentische Bewertungen dar, bringt aber möglicherweise eine geringere Aussagekraft für die Verbraucher mit sich.

Bei den Geschäftsmodellen anderer Portale besteht keine Möglichkeit, ein geschlossenes System zu betreiben; sie verwenden ein offenes Bewertungssystem. Jameda z. B. bietet erst in jüngerer Zeit und in sehr eingeschränktem Maß die Vermittlung von Behandlungsterminen an. Das Geschäftsmodell basiert aber darauf, dass möglichst viele Patienten von ihrem Arztbesuch berichten und Verbraucher sich deshalb auf dem Portal umfassend informieren können, auch wenn der Kontakt mit der Arztpraxis ganz überwiegend direkt erfolgt, also keine „Nutzungskontrolle“ durch Jameda stattfinden kann. Auch Anbieter wie Tripadvisor sind deshalb attraktiv für Nutzer, weil sie eine sehr große Zahl an Bewertung für sehr viele Restaurants u. a. zeigen. Die Vermittlung von Reservierungen spielt dagegen nur eine sehr untergeordnete Rolle, so dass auch Tripadvisor für sein Geschäftsmodell auf die Verwendung eines offenen Bewertungssystems angewiesen ist.

Einige Portale sind bereits in der Vergangenheit gegen von ihnen identifizierte Vermittler nicht-authentischer Bewertungen auf ihrem Portal zivilrechtlich vorgegangen; von den befragten Portalen haben dies Holidaycheck, Jameda und Mobile.de berichtet (s. Kapitel E.3.1.). Holidaycheck konnte dabei die mangelnde Authentizität der in Frage stehenden Bewertungen eindrucksvoll nachweisen: Die durch den beklagten Bewertungsvermittler zustande gekommenen Bewertungen wiesen ausnahmslos sechs von sechs möglichen Sonnen auf. Außerdem hatte der Vermittler angeboten, dass den Bewertenden Texte und Fotos von den Anbietern zur Verfügung gestellt werden, die dann in die Bewertung übernommen würden. Gerade die Fotos können dazu dienen, dass ein Verbraucher gar nicht auf die Idee kommt, dass es sich um eine nicht-authentische Bewertung handelt. Das LG Berlin untersagte in seiner einstweiligen Verfügung dazu die Veröffentlichung bzw. das Veranlassen der Veröffentlichung von Bewertungen. Das Gericht sah als erwiesen an, dass die Verfasser der Bewertungen überhaupt nicht Gast in dem mit Bestnote bewerteten Beherbergungsbetrieb gewesen waren.¹²⁵ Auch Mobile.de verwies in einem selbst angestregten Verfahren darauf, dass der Bewertungsvermittler auf seiner Seite unter anderem folgendes angeboten habe: „Gerne können Sie uns [...] Textvorschläge, Bilder oder einzelne Bauteile, auf die besonders eingegangen werden soll, zukommen lassen. Diese werden wir dann den Rezensenten zur Verfügung stellen.“ Nichts deutet dagegen darauf hin, dass die Verfasser der Bewertungen in irgendeiner Form dazu aufgefordert oder angehalten wurden, den jeweiligen Autohändler tatsächlich aufzusuchen und einen Testkauf durchzuführen. Das LG Hamburg sah im Angebot eines Bewertungsvermittlers, auf Mobile.de tätigen Händlern positive und, wie sich aus den Entscheidungsgründen ergibt, nicht-authentische Bewertungen zu vermitteln, eine unlautere geschäftliche Handlung im Sinne von § 3 Abs. 1 UWG. Das Angebot sei darauf ausgerichtet, Verbraucher irrezuführen.¹²⁶

¹²⁵ LG Berlin, Beschl. v. 20.11.2018, Az. 15 O 203/18 (nicht veröffentlicht).

¹²⁶ LG Hamburg, Urt. v. 07.03.2019, Az. 403 HKO 3/19 (nicht veröffentlicht).

F.4.2. Lösungsansatz: Stärkere Inpflichtnahme der Portale

Trotz gefestigter Rechtsprechung, nicht-authentische Bewertungen als Verbraucherrechtsverstoß einzuordnen, findet die Vermittlung und Veröffentlichung nicht-authentischer Bewertungen weiter statt. Die bestehenden Urteile in dem Bereich zeigen allerdings, dass der geltende Rechtsrahmen grundsätzlich ausreicht, um nicht-authentische Bewertungen als verbraucherschädigende Irreführung zu ahnden. Der Grund für die mangelhafte Verfolgung solcher Bewertungen dürfte daher wie bei den verdeckten Produkttest-Bewertungen und den manipulierten Bewertungen eher im Bereich der *Rechtsdurchsetzung* liegen. Portale, auf denen nicht-authentische Bewertungen gehäuft auftreten, sind bisher nicht dazu verpflichtet worden, diese im Rahmen zumutbarer spezifischer Überwachungspflichten proaktiv zu bekämpfen. Einige auf diesem Gebiet bereits aktiven Portale haben allerdings auch auf Schwierigkeiten hingewiesen, die ihnen in Deutschland rechtlich bei der Bekämpfung nicht-authentischer Bewertungen begegnen, und wünschen sich an dieser Stelle mehr Unterstützung (s. u.).

Portale sollten stärker in die Pflicht genommen werden, nicht-authentische Bewertungen auf ihren Seiten zu identifizieren und zu entfernen. Dies sieht auch der Gesetzgeber so, denn auch die neue Nr. 23b des Anhang I der UGP-Richtlinie, die derzeit in nationales Recht umgesetzt wird, geht in diese Richtung. Danach stellt „die Behauptung, dass Bewertungen eines Produkts von Verbrauchern stammen, die das Produkt tatsächlich verwendet oder erworben haben, ohne dass angemessene und verhältnismäßige Schritte unternommen wurden, um zu prüfen, ob die Bewertungen wirklich von solchen Verbrauchern stammen“ eine irreführende Geschäftspraktik dar. Die Vorschrift könnte sich v. a. gegen Portale wenden, die in fälschungsanfälligen Branchen ein offenes Bewertungssystem betreiben, ohne standardisiert Filtermaßnahmen zur Überprüfung der Authentizität von Bewertungen durchzuführen. Dennoch erscheint die Praxis einiger Portale auch schon mit dem derzeitigen Rechtsrahmen nicht vereinbar zu sein: Die Ermittlungen im Rahmen der Sektoruntersuchung haben gezeigt, dass einige Portale bezogen auf nicht-authentische Bewertungen überhaupt keine Maßnahmen ergreifen, obwohl bekannt sein dürfte, dass gerade offene Bewertungssysteme der Gefahr nicht-authentischer Bewertungen in beträchtlichem Umfang ausgesetzt sind und obwohl die Vorgehensweise einzelner Portale zeigt, dass fehlende Authentizität mittels automatischer und manueller Auswertung von Metadaten zumindest teilweise aufgespürt werden kann. Solange die Versäumnisse nicht offen gelegt werden und damit ein entsprechender Imageschaden für das Portal entsteht, haben viele Portale bisher offenbar keinen Anreiz, viel in die Aufdeckung nicht-authentischer Bewertungen zu investieren. Bisher ist auch kein Fall bekannt, in dem ein Portal dafür verklagt wurde, dass es systematisch zu wenig gegen nicht-authentische Bewertungen unternimmt. Wenn es zu zivilrechtlichen Verfahren kam, waren es häufig die Portale selbst, die dabei als Kläger gegenüber Anbietern oder Bewertungsvermittlern aufgetreten sind. Dabei wäre es durchaus denkbar, die Portale, die sich bisher noch nicht aktiv um die Bekämpfung nicht-authentischer Bewertungen bemühen, dafür zur Verantwortung zu ziehen. Gerade wenn die Rahmenbedingungen auf einem Portal so ausgestaltet sind, dass nicht-authentische Bewertungen wahrscheinlich sind, unterliegt ein Portal proaktiven Überwachungspflichten, um entsprechende Rechtsverstöße aufzudecken. Derartige Rahmenbedingungen sind z. B. ein offenes Bewertungssystem oder ein hoher Einfluss von Anzahl und Güte der Bewertungen auf die Auffindbarkeit eines Items. Als Kläger kommen dafür z. B. Verbraucherschutzorganisationen in Frage.

Gegebenenfalls benötigen Portale allerdings auch mehr rechtliche Möglichkeiten bei der Bekämpfung nicht-authentischer Bewertungen. Die Ermittlungen haben gezeigt, dass einige große Portale mit offenen Bewertungssystemen sehr umfangreich versuchen, nicht-authentische Bewertungen zu identifizieren, zu löschen bzw. gar nicht erst zu veröffentlichen. Diese Portale schaffen es aufgrund ihrer Filtermaßnahmen offenbar, intern zumindest einen gewissen Anteil nicht-authentischer Bewertungen zu erkennen. Das Vorgehen ist dabei sehr differenziert und hängt stark davon ab, in welcher Branche ein Portal tätig ist: Maßnahmen im Hotelbereich sind nicht ohne weiteres auf Ärztebewertungen übertragbar und umgekehrt. Eine wichtige Rolle spielen in allen Bereichen Algorithmen, die eingehende Bewertungen nach einer Vielzahl von Kriterien mit vorhandenen Bewertungen vergleichen und Auffälligkeiten dem Qualitätsmanagement des Portals melden, das dann weitere Schritte einleitet. Jameda hat berichtet, dass Auffälligkeiten überprüft würden, indem z. B. eine zusätzliche Identifikation des Verfassers der Bewertung per SMS verlangt wird oder indem gezielt Verfasser von Bewertungen angeschrieben werden, um zu klären, ob sie den bewerteten Arzt tatsächlich aufgesucht haben bzw. ihre Bewertung unbeeinflusst zustanden gekommen ist.

Trotz dieser Maßnahmen gelingt es aber auch den engagierten Portalen offenbar nicht, die Entstehung neuer nicht-authentischer Bewertungen zu verhindern. Ein Portal hat darauf hingewiesen, dass es aufgrund der geltenden Rechtslage in Deutschland schwierig sei, nicht-authentische Bewertungen nachhaltig einzuschränken. Insbesondere komme in Deutschland eine strafrechtliche Verfolgung von Bewertungsvermittlern, wie sie in Italien oder Großbritannien möglich sei, nicht in Frage. Bei der zivilrechtlichen Verfolgung dieser Unternehmen nach den Vorschriften des UWG müsse zunächst das Portal als Mitbewerber nach § 4 UWG angesehen werden. Schon daran könne eine Verfolgung scheitern. Auch die Erreichbarkeit von Bewertungsvermittlern im Ausland wurde als Problem geschildert. Ein Weg, Bewertungsvermittler zu identifizieren, könnte deren deutsche Bankverbindung sein. Hier bestehe aber keine Auskunftspflicht gegenüber einem Portalbetreiber.

F.5. Asymmetrisch selektierte Bewertungen

Das Problem der asymmetrisch selektierten Bewertungen ist dadurch gekennzeichnet, dass zwar die einzelnen Bewertungen authentisch und manipulationsfrei sind, aber die Gesamtheit der auf einem Portal dargestellten Bewertungen dadurch verzerrt wird, dass eine bestimmte Art abgegebener Bewertungen – in der Regel negative – mit größerer Wahrscheinlichkeit nicht dargestellt wird. Die Ermittlungen haben im Wesentlichen zwei Quellen für solche Verzerrungen ergeben. Zum einen gibt es im Bewertungssystem selbst verankerte Asymmetrien: Teilweise werden Anstupser oder Anreize zur Bewertungsabgabe verwendet, die asymmetrisch wirken, weil dem Bewertungsverfasser z. B. die Abgabe einer positiven Bewertung leichter gemacht wird (nur ein Klick) als die Abgabe einer negativen Bewertung (Begründung erforderlich) oder weil z. B. Gutscheine oder Rabatte nur im Fall einer positiven Bewertung in Aussicht gestellt werden. Teilweise werden auch Filter so eingesetzt, dass negative eher als positive Bewertungen davon erfasst werden, z. B. wenn Bewertungen mit weniger als drei Sternen immer zusätzlich manuell überprüft werden und damit eine höhere Löschungswahrscheinlichkeit besteht. Zum anderen führt das Notice-and-take-down-Verfahren, bei dem ein

Portal (nur) nachträglich von Dritten als kritisch gemeldete Bewertungen kontrolliert und ggfs. gelöscht werden, ebenfalls dazu, dass der Anteil der dargestellten negativen Bewertungen sinkt bzw. der Anteil positiver Bewertungen steigt.

F.5.1. Ökonomische und rechtliche Bewertung

Die Ermittlungen haben ergeben, dass Portale und/oder Anbieter in Einzelfällen asymmetrisch wirkende Anstupsen verwenden, um gezielt nur zufriedene Kunden zur Abgabe einer Bewertung zu motivieren (s. Kapitel D.1.3.). Neben Aufforderungen, die gezielt zufriedene Kunden ansprechen, gehören dazu auch alle Arten asymmetrischer Bewertungsabläufe, die die Abgabe einer negativen Bewertung für den Nutzer umständlicher oder langwieriger machen als die Abgabe einer positiven Bewertung. Dies führt dazu, dass der Anteil der abgegebenen positiven Bewertungen höher ist als bei einem symmetrischen Bewertungsablauf. Außerdem verwenden einzelne Portale Filtermechanismen, die dafür sorgen, dass negative Bewertungen mit höherer Wahrscheinlichkeit herausgefiltert und deshalb nicht angezeigt werden als positive. Das ist z. B. der Fall, wenn alle Bewertungen mit einem oder zwei Sternen vor ihrer Veröffentlichung grundsätzlich manuell geprüft werden, alle anderen aber nur im Fall einer Markierung durch einen Wortfilter (s. Kapitel D.2.1.). Diese asymmetrischen Komponenten im Bewertungssystem führen dazu, dass der Leser relativ mehr positive Bewertungen sieht und so einen positiv verzerrten Eindruck von der Beschaffenheit des bewerteten Items bekommt. Da die Anzahl und Breite insgesamt dargestellter Bewertungen durch die asymmetrische Behandlungsweise nicht steigt, sondern sogar reduziert wird, ist im Gegensatz zu incentivierten Bewertungen auch kein ökonomischer Vorteil eines solchen Vorgehens zu erkennen, der gegen die Nachteile der Verzerrung des Wettbewerbs und der Irreführung des Verbrauchers aufgewogen werden könnte.

Asymmetrische Komponenten eines Bewertungssystems stellen außerdem eine verbraucherrechtliche Irreführung z. B. jedenfalls nach § 5 Abs. 1 Nr. 1 UWG dar, für die das jeweilige Portal als Betreiber des Bewertungssystems auch verantwortlich ist. Das gilt sowohl für asymmetrische Anstupsen oder Anreize zur Bewertungsabgabe als auch für die asymmetrischen Komponenten im Filtermechanismus, durch die jeweils zur Täuschung geeignete Angaben über die Beschaffenheit der angebotenen Items sowie ggfs. ihre Art, Ausführung, Vorteile und Risiken entstehen: Die Gesamtheit der Bewertungen ist positiv verzerrt, da systembedingt zufriedene Kunden eher eine Bewertung abgeben als unzufriedene bzw. die Bewertungen zufriedener Kunde mit größerer Wahrscheinlichkeit auch tatsächlich dargestellt und nicht (heraus)gefiltert werden. Die Beschaffenheit der angebotenen Items wird dadurch zu positiv dargestellt, Vorteile werden ggfs. besonders stark betont, wohingegen Risiken weniger häufig oder sogar gar nicht erwähnt werden. Die geschäftliche Handlung ist deshalb irreführend.

Bezogen auf eine durch Filtermaßnahmen bewirkte Asymmetrie ist der Tatbestand der Irreführung bereits durch zwei Urteile des Oberlandesgerichts Düsseldorf bzw. des Bundesgerichtshofs bestätigt worden. Die beiden Urteile betreffen zwar unabhängige Sachverhalte, beziehen sich aber jeweils auf die gleiche Praxis desselben Bewertungsportals. Dieses Portal veröffentlichte in einem gewissen Zeitraum positive Bewertungen direkt nach Abgabe, wohingegen negative Bewertungen vor der Veröffentlichung auf Rechtswidrigkeit

geprüft sowie einem Schlichtungsverfahren unterworfen und erst danach – je nach Ausgang des Verfahrens – veröffentlicht wurden.

In dem vom OLG Düsseldorf zu entscheidenden Fall¹²⁷ ging es um ein Zahnersatzprodukt und damit ein Heilmittel. Werbliche Maßnahmen für Heilmittel sind spezialgesetzlich durch das HWG geregelt. Zur asymmetrischen Selektion im Filterprozess führt das OLG aus, dass das Bewertungssystem verschiedene Komponenten beinhalte, die zu einer asymmetrischen Selektion und damit zu einer Fälschung des Ergebnisses beitragen: So würden positive Bewertungen sofort, negative erst nach fünf Tagen eingestellt, da der Betreiber des Bewertungssystems dem Anbieter zunächst ein Schlichtungsverfahren vorschlägt. Zwar nehme die Bedeutung dieses statistischen Fehlers mit fortschreitender Bewertungsdauer ab, das Schlichtungsverfahren wirke sich aber noch in einer weiteren, weitaus gravierenderen Weise auf das Ergebnis aus: Allein seine Existenz halte einen Teil der unzufriedenen Kunden von der Abgabe einer negativen Bewertung ab. Viele Menschen seien konfliktscheu und daher versucht, allein wegen der Möglichkeit, ihre negative Bewertung verteidigen zu müssen, von der Abgabe einer Bewertung abzusehen. Schließlich würden negative Bewertungen grundsätzlich auf rechtswidrige Inhalte geprüft, positive nicht.¹²⁸ In allen drei Punkten sieht das Gericht einen Verstoß gegen § 11 Abs. 1 Nr. 11 HWG. Demnach darf außerhalb der Fachkreise für das Heilmittel u. a. dann nicht mit Äußerungen Dritter geworben werden, wenn diese in irreführender Weise erfolgen. Das sah das Gericht aufgrund der asymmetrischen Selektion hier als erfüllt an. Eine Prüfung im Hinblick auf § 5 UWG hat das Gericht in diesem Fall aufgrund der Spezialität des § 11 Abs. 1 Nr. 11 HWG nicht vorgenommen.¹²⁹ Da sich die Argumentation des Gerichts jedoch auf das Vorliegen einer „Irreführung“ bezieht und das HWG ähnlich dem UWG das Publikum vor unrichtiger bzw. unsachlicher Werbung schützen will, lässt sich die Entscheidung des Gerichts aber durchaus auf Bewertungen von Items, die nicht dem HWG unterfallen, übertragen.

Der Bundesgerichtshof bejaht bei einer Prüfung der Filtermaßnahmen desselben Bewertungssystems, im konkreten Fall allerdings für einen Anbieter von Druckerzeugnissen, ebenfalls eine Irreführung, wenn vom Betreiber des Bewertungssystems nicht sofort alle Bewertungen ungefiltert eingestellt werden. Wenn wie im zu entscheidenden Fall mit „garantiert echten Meinungen“ geworben werde, gehe der Verbraucher davon aus, dass positive und negative Bewertungen gleichermaßen ungefiltert dargestellt und in die Ermittlung der durchschnittlichen Kundenmeinung eingehen würden. Zur Vermeidung einer Irreführung bei einer Werbung mit den Kundenbewertungen müsse deutlich über das Schlichtungsverfahren aufgeklärt werden.¹³⁰

Auch der Notice-and-take-down-Prozess, den einige der befragten Portale durchführen, wenn bewertete Anbieter sie zu einer nachträglichen Überprüfung auffordern, führt zu einer asymmetrischen Selektion von Bewertungen (s. Kapitel D.2.4.), da die Aufforderung zur nachträglichen Überprüfung in der Regel für negative Bewertungen erfolgt. Bereits während der Prüfung der einzelnen (negativen) Bewertung werden diese in der Regel nicht mehr angezeigt; ein nicht geringer Teil wird am Ende des Prüfprozesses auch endgültig entfernt.

¹²⁷ OLG Düsseldorf, Urt. v. 19.02.2013, Az. I-20 U 55/12 – *Kundenbewertungen*.

¹²⁸ OLG Düsseldorf, Urt. v. 19.02.2013, Az. I-20 U 55/12 – *Kundenbewertungen* = juris Rn. 26, 26.

¹²⁹ OLG Düsseldorf, Urt. v. 19.02.2013, Az. I-20 U 55/12 – *Kundenbewertungen* = juris Rn. 27.

¹³⁰ BGH, Urt. v. 21.01.2016, Az. I ZR 252/14, Rn. 38 – *Kundenbewertung im Internet*.

Positive Bewertungen werden dagegen naturgemäß von den Bewerteten eher nicht kritisiert und deshalb auch im Rahmen dieses Prozesses nicht überprüft.

In den letzten Jahren sind zahlreiche Anbieter, vor allem Ärzte und Zahnärzte, gegen für sie unliebsame negative und vermeintlich nicht authentische Bewertungen zivilrechtlich vorgegangen. Daraus hat sich eine umfassende Rechtsprechung entwickelt, die den Portalen relativ umfangreiche Prüfpflichten auferlegt, die reaktiv auf die Kenntnisnahme eines möglichen Verstoßes zu erfüllen sind. Proaktive Überwachungspflichten der Portale sind daraus zumindest bisher allerdings kaum abgeleitet worden.

Aus ökonomischer Sicht hat diese Rechtsprechung nicht unerheblichen Schaden angerichtet, da sie dazu führt, dass auf den von derartigen nachträglichen Beschwerden besonders betroffenen Portalen eine große Zahl negativer Bewertungen, die hilfreich für Verbraucher wären, entfernt werden, was auch als „Overblocking“ bezeichnet wird.¹³¹

Es ist davon auszugehen, dass bei der nachträglichen Überprüfung im Rahmen des Notice-and-take-down zu viele Bewertungen gelöscht werden, also auch solche, die authentisch und ansonsten ordnungsgemäß zustanden gekommen sind. Die Gründe dafür liegen im Prozess selbst und den Risiken für das Portal (s. Kapitel D.2.4.). Viele Bewertungsverfasser reagieren nicht auf eine Aufforderung des Portals, Authentizitätsnachweise zu erbringen. Dem Portal bleibt dann nichts anderes übrig, als eine Bewertung zu löschen. Außerdem ist zu erwarten, dass ein Portal im Zweifelsfall lieber zu viele als zu wenige Bewertungen löscht, um drohende Rechtsstreitigkeiten zu vermeiden.

Dazu kommt, dass den Anbietern die Situation der Bewertungsportale bewusst ist und deshalb ihrerseits ein starker Anreiz besteht, bei jeder unliebsamen Bewertung zumindest zu versuchen, diese entfernen zu lassen. Mehrere Portale berichteten von massenhaften, standardisiert und oberflächlich verfassten Anwaltsschreiben, die stets die Authentizität unliebsamer Bewertungen in Frage stellten.

F.5.2. Lösungsansatz: Beseitigung asymmetrischer Komponenten

Unabhängig von den Details des jeweiligen Falles bestätigen die beiden Urteile, dass Maßnahmen des Betreibers eines Bewertungssystems, die dazu führen, dass negative Bewertungen in geringerem Maße abgegeben bzw. angezeigt werden als positive Bewertungen, eine Irreführung des Verbrauchers darstellen. In diesen Konstellationen ist auch nicht erkennbar, dass das Problem der Irreführung durch eine entsprechende Aufklärung der Verbraucher behoben werden kann. Insofern besteht die einzige Lösung des Problems einer asymmetrischen Selektion von Bewertungen darin, die Komponenten des Bewertungssystems, die zur Asymmetrie führen, zu entfernen. Das sollte in der Regel aber kein Problem sein.

Verantwortlich sind dafür in den meisten Fällen die Portale, die asymmetrische Anstupser oder Anreize setzen bzw. asymmetrisch filtern. Asymmetrische Anstupser oder Anreize können in Einzelfällen auch von einem

¹³¹ *Specht*, Manipulierte Kundenbewertungen auf Online-Vermittlungsportalen – eine lauterkeitsrechtliche Betrachtung, in: Rott/Tonner (Hrsg.): Online-Vermittlungsplattformen in der Rechtspraxis, 2018, S. 81-93.

Anbieter auf einer Plattform nur für seine eigenen Leistungen gesetzt werden. In diesen Fällen könnte erwogen werden, ob das betroffene Portal ebenfalls eine mittelbare Verantwortlichkeit dafür trägt. Wie für den Fall incentivierter Bewertungen scheint dies allerdings auch hier nicht notwendig zu sein, um einen effektiven Rechtsschutz zu garantieren: Der eigentliche Verursacher des Rechtsverstoßes, nämlich der Anbieter, der den asymmetrischen Anstupser oder Anreiz setzt, ist in der Regel öffentlich bekannt und kann daher selbst zur Verantwortung gezogen werden.

Schwieriger als bei systemisch bedingter Asymmetrie ist der Umgang mit Asymmetrie, die durch den Notice-and-take-down-Prozess entsteht. Asymmetrische Filter- und Anreizkomponenten können entfernt werden. Eine Aufgabe des Notice-and-take-down-Prozesses kommt dagegen nicht in Frage: denn Portale haben nicht die Möglichkeit, sich diesem Prozess zu entziehen, der – gerichtlich bestätigt – dem legitimen Ausgleich zwischen dem Recht des Betroffenen auf den Schutz seiner Persönlichkeit und dem Recht von Portal bzw. Verfasser auf Meinungs- und Medienfreiheit dient.¹³² Nach den Ermittlungen sind nur wenige Portale massenhaft von derartigen Verfahren betroffen. Auf diesen Portalen kann eine großflächige Entfernung oder das Nicht-Einstellen negativer unliebsamer Bewertungen das Gesamtbild, das ein Verbraucher bekommt, spürbar verzerren. Portale könnten die Verbraucher durch einen entsprechenden Hinweis auf diese Gefahr aufmerksam machen.

F.6. Weitere Problembereiche

Neben den genannten Problemsachverhalten wurden im Laufe der Sektoruntersuchung weitere Problembereiche identifiziert. So hat sich in einem Fall ergeben, dass die vollständige Aufklärung und Bewertung der angezeigten Kundenäußerungen nur in einem Verwaltungsverfahren gegen das Unternehmen möglich gewesen wäre, was dem Bundeskartellamt im Verbraucherschutzbereich jedoch nicht möglich ist. Die Sektoruntersuchung hat außerdem gezeigt, dass an zwei weiteren Stellen Verbesserungspotential bezogen auf Bewertungen besteht. Das betrifft die Angabe von Durchschnittsbewertungen bei einer geringen Anzahl von Bewertungen für ein einzelnes Item und – wesentlich wichtiger – die geringe Kenntnis vieler Portale darüber, wie repräsentativ die Meinungen der Verbraucher, die sich in den Bewertungen spiegeln, sind. Durch eine gesteigerte Repräsentativität von Bewertungen könnte der Nutzen, den Verbraucher aus Bewertungen ziehen, möglicherweise noch deutlich gesteigert werden.

F.6.1. Unklare Herkunft und Auswahl von Kundenbewertungen

Bei einem der befragten Portale, Pearl, konnte nicht in einer für eine abschließende Beurteilung hinreichenden Weise geklärt werden, woher die dargestellten „Kundenmeinungen“ bzw. „Kundenbeiträge“ tatsächlich stammen sowie ob und in welcher Weise eine Filterung erfolgt. So ist das Unternehmen dem ursprünglichen Verdacht des Amtes, es handele sich um ein fiktives Bewertungssystem, zwar entgegengetreten, trug dabei aber nicht widerspruchsfrei vor und ist Belege schuldig geblieben. Dies weiterzuverfolgen, hätte den Rahmen

¹³² BGH, Urt. v. 01.03.2016, Az. VI ZR 34/15 – *jameda.de II*.

einer Sektoruntersuchung überschritten, die in überschaubarer Zeit Ergebnisse in Bezug auf einen gesamten Wirtschaftszweig erbringen soll. Während bei Kartellrechtsverstößen in solchen Fällen im Anschluss ein Verwaltungsverfahren gegen das betroffene Unternehmen eingeleitet würde, steht dies dem Bundeskartellamt im Verbraucherschutzbereich von Gesetzes wegen nicht offen.

F.6.2. Missverständliche Durchschnittsangaben

Eigentlich alle Portale, die Nutzerbewertungen anzeigen, stellen auch eine Zusammenfassung dieser Bewertungen dar (s. Kapitel D.3.2.), in der Regel handelt es sich dabei um den arithmetischen Mittelwert, teilweise wird auch die Verteilung der Sternebewertungen angezeigt. Verbunden damit wird in der Regel angezeigt, wie viele Bewertungen das entsprechende Item hat. Dieser Angabe kommt eine hohe Bedeutung zu, da sie an prominenter Stelle platziert ist und von den Verbrauchern in der Regel zuerst wahrgenommen wird. Es ist davon auszugehen, dass die Angabe auch Verbraucher beeinflusst, die wenig Wert auf den Inhalt von Bewertungen legen und nicht explizit einzelne Rezensionen lesen. Durchschnittsangaben sind auch grundsätzlich zulässig und führen nach geltender Rechtsprechung nicht dazu, dass das Portal sich die zugrunde liegenden Bewertungen zu eigen macht und damit Verantwortung für deren Inhalte übernehmen muss.¹³³

Wenn für ein Item nur sehr wenige Bewertungen vorliegen, kann dies zu Schwierigkeiten bei der Durchschnittsbewertung führen. Der Durchschnitt kann durch Ausreißer verzerrt sein bzw. hat nur eine eingeschränkte Aussagekraft. Teilweise zeigen Portale daher Durchschnittswerte erst ab einer bestimmten Zahl von Bewertungen an. Dieses Vorgehen kann Items mit wenigen Bewertungen benachteiligen, weil Verbraucher eher von Items mit Durchschnittsangabe angezogen werden, allein wegen der Farbigkeit der Sterne. Andere Portale ergänzen die tatsächlichen Bewertungen durch fiktive Bewertungen in durchschnittlicher Höhe, um einen gemäßigten Durchschnittswert zu erhalten. Dies kann den Verbraucher allerdings irreführen, der davon ausgeht, dass der Durchschnittswert nur auf tatsächlichen Bewertungen beruht.

Es gibt verschiedene Möglichkeiten, das Problem von Ausreißern und mangelnder Aussagekraft bei einer geringen Zahl von Bewertungen anzugehen. Wichtig ist dabei vor allem, dass für den Verbraucher transparent ist, welche Regeln den Angaben zugrunde liegen.

Das in den vorhergehenden Abschnitten dargestellte Problem der asymmetrischen Verzerrung von Bewertungen spielt auch für die Durchschnittsbewertung eine Rolle, da hier lediglich die tatsächlich angezeigten Bewertungen einbezogen werden.

F.6.3. Fehlende Repräsentativität

Ein wenig beachtetes Problem von Bewertungen liegt schließlich darin, dass sie häufig ein wenig repräsentatives Bild der Meinung der Gesamtheit aller Nutzer zu einem Item abgeben. Der Anteil der Verbraucher, der

¹³³ BGH, Urt. v. 19.03.2015, Az. I ZR 94/13, Rn. 28 – *Hotelbewertungsportal*.

Bewertungen abgibt, ist äußerst gering (s. Kapitel E.1.2.). Es handelt sich dabei zudem nicht um einen Querschnitt aller Nutzer. Vielmehr sind es eher die besonders Zufriedenen oder Unzufriedenen, die bewerten, und bestimmte Verbrauchergruppen bewerten mit höherer Wahrscheinlichkeit als andere. Die Sektoruntersuchung hat ergeben, dass sich die meisten Portale wenige oder überhaupt keine Gedanken darüber machen, ob die Bewertungen für Items auf ihrer Seite ein repräsentatives Bild ergeben. Darüber hinaus werden von vielen Portalen erst recht keine Maßnahmen ergriffen, mögliche Verzerrungen zu korrigieren.

Nutzerbewertungen stellen für viele Verbraucher im Internet eine wichtige Entscheidungshilfe dar. Der Wert der Informationen, den Verbraucher aus Nutzerbewertungen erhalten, könnte noch wesentlich höher sein, wenn eine gewisse Repräsentativität der Bewertungen gewährleistet werden könnte. Es wäre daher zu begrüßen, wenn Portale und Anbieter in diesem Bereich mehr Initiative zeigen und Verbraucherschützer, Politik und Medien dieses Problem stärker berücksichtigen würden. Dazu müsste zunächst untersucht werden, ob und ggfs. welche Verbrauchermeinungen sich nicht in den Nutzerbewertungen widerspiegeln. Anschließend könnten Portale und Anbieter z. B. versuchen, durch gezielte Ansprache genau diese Gruppen zur Abgabe von Bewertungen anzuregen.

G. Fazit

Nutzerbewertungen im Internet stellen eine wichtige Informationsquelle für Verbraucher dar. Für viele Items gibt es aber zu wenige Bewertungen, denn nur selten sind Nutzer bereit, ohne Gegenleistung Bewertungen zu verfassen. Gleichzeitig sind Bewertungen aber für die Anbieter von großer Bedeutung, um von Verbrauchern ausgewählt bzw. überhaupt gefunden zu werden. Daher kommt es regelmäßig zu Fälschungen und Manipulationen. Die von den Portalen durchgeführten Maßnahmen zur Bekämpfung solcher unzulässiger Bewertungen sind nicht immer ausreichend. Einige Portale müssen mehr unternehmen, um derartige Bewertungen zu identifizieren. Darüber hinaus sollten Portale bei Bedarf Möglichkeiten schaffen, über monetäre Anreize zur Bewertungsabgabe wie kleine Belohnungen, Gutscheine oder Gewinnspiele oder den Einsatz bzw. die Zulassung von Produkttestern zusätzliche Bewertungen zu generieren. Anreize und Produkttests müssen allerdings für Verbraucher klar als solche erkennbar sein. Mit der vorliegenden Sektoruntersuchung schafft das Bundeskartellamt ein hohes Maß an Aufklärung und Transparenz zum Thema Nutzerbewertungen und gibt gleichzeitig konkrete Hinweise für Verbesserungen des Verbraucherschutzes in diesem Bereich.

Die aktuelle Diskussion über „Fake-Bewertungen“ macht deutlich, dass die Zuverlässigkeit und die Vertrauenswürdigkeit von Nutzerbewertungen als wichtiges Informations- und Wettbewerbsinstrument gefährdet sind. Die Ermittlungen des Bundeskartellamtes im Rahmen der Sektoruntersuchung haben gezeigt, dass die damit verbundenen Probleme vielschichtig und komplex sind und eine differenzierte Betrachtung erfordern. So weisen die von den Portalen betriebenen Bewertungssysteme unterschiedliche Defizite auf, die einzelnen Marktteilnehmer verfolgen verschiedene Interessen und die identifizierten Problemkategorien sind jeweils individuell zu bewerten.

Insbesondere auf Online-Handelsplattformen stellen verdeckte Produkttests ein Problem dar, da diese auch für aufmerksame Verbraucher nicht als solche zu erkennen sind. Häufig sind diese Bewertungen außerdem inhaltlich manipuliert, so dass der Verbraucher ggfs. viele sehr gute Bewertungen für ein eigentlich mittelmäßiges Produkt findet. Eine Ursache für das massenhafte Auftreten von Produkttest-Bewertungen ist die hohe Bedeutung der Bewertungen für Anbieter bei der gleichzeitig geringen Bereitschaft der Verbraucher, regelmäßig Bewertungen zu schreiben. Die dadurch entstehende Lücke zwischen Angebot und Nachfrage von Bewertungen sollte allerdings nicht durch verdeckte und ggfs. manipulierte Produkttest-Bewertungen geschlossen werden, sondern durch offene und unbeeinflusste Anreize zur Bewertungsabgabe oder den Einsatz von Produkttestern. Das Zustandekommen solcher Bewertungen ist verbraucherrechtskonform, sofern es dem Verbraucher gegenüber transparent gemacht wird. Betroffene Portale können mit vertretbarem Aufwand die Rahmenbedingungen dafür schaffen, dass die Abgabe und Kennzeichnung solcher Bewertungen auf ihrer Seite möglich ist.

Die Vorab-Filterung von Bewertungen, die Portale durchführen, fokussiert sich bisher bei vielen Portalen auf Hate Speech, Diskriminierungen und Datenschutzverstöße. Die Identifizierung von nicht authentischen oder manipulierten Bewertungen fällt dagegen deutlich zurück. Einige Ausnahmen von Portalen zeigen allerdings, dass mit vertretbarem Aufwand viel mehr möglich ist: Diesen Portalen gelingt es, durch eine branchenspezifisch zugeschnittene Prüfung von Bewertungen zahlreiche Verdachtsfälle zu identifizieren, die dann vom Qualitätsmanagement des jeweiligen Portals weiterverfolgt werden.

In einigen Bereichen hat das Bundeskartellamt festgestellt, dass viel mehr positive als negative Bewertungen dargestellt werden, entweder weil die Abgabe negativer Bewertungen erschwert wird oder weil im Verlauf des Filterprozesses und der nachträglichen Kontrolle aufgrund von Beschwerden systematisch mehr negative als positive Bewertungen gelöscht werden. Hier können und sollten Portale darauf achten, dass solche Verzerrungen möglichst vermieden oder zumindest die Verbraucher darüber aufgeklärt werden.

Die Sektoruntersuchung hat einige Missstände bei Nutzerbewertungen aufgedeckt. Da das Bundeskartellamt im Verbraucherschutz zwar über Aufklärungs-, nicht aber über Durchsetzungsbefugnisse verfügt, kann es selbst in einem konkreten Verdachtsfall, wie er sich hier ergeben hat, kein Verwaltungsverfahren gegen das betreffende Unternehmen führen. Hier und auch im Übrigen müssen Missstände von den Marktteilnehmern selbst oder über den Zivilrechtsweg behoben werden. Insbesondere die Portale als die Betreiber der Bewertungssysteme hätten die Möglichkeit, durch Anpassungen ihrer Bewertungssysteme mehr verbraucherrechtskonforme Bewertungen zu ermöglichen und einige der dargestellten Missstände zu beheben, um das Vertrauen der Verbraucher in Nutzerbewertungen auch dauerhaft zu erhalten. Darüber hinaus sind private Akteure und Verbraucherschutzorganisationen gefragt, ein rechtskonformes Verhalten der Marktteilnehmer einzufordern, soweit sie dazu trotz ihrer begrenzten Nachweismöglichkeiten in der Lage sind. Im Rahmen von Rechtstreitigkeiten sollte dabei auch auf eine klarere und stärkere Verantwortlichkeit von Portalen für die auf ihren Webseiten begangenen Verbraucherverstöße geachtet werden.

H. Anhang

Befragte Portale

Portal	Branche (Schwerpunkt)	Antworten erhalten zu den Fragebögen		
		Produkt- bewertungen	Unternehmens- bewertungen	App- Bewertungen
123Recht.de/ Frag-einen-anwalt.de	Anwälte/Finanzen		x	
Ab-in-den-urlaub.de	Reisen		x	
Airbnb.de	Reisen		x	
Alternate.com	Elektronikartikel	x		
Amazon.de	Generalist	x	x	
Anwalt.de	Anwälte/Finanzen		x	
Apple Appstore	Apps			x
Autoscout24.de	Autohändler		x	
Bewertet.de	Handwerker		x	
Blauarbeit.de	Handwerker		x	
Booking.com	Reisen		x	
Check24.de	Generalist		x	
Conrad.de	Elektronikartikel	x		
Dasoertliche.de	Generalist		x	
Dastelefonbuch.de	Generalist		x	
Dm.de	Generalist	x		
Docinsider.de	Ärzte		x	
Docmorris.de	Medikamente	x		
Ebay.de	Generalist	x	x	
Ebay-kleinanzeigen.de	Generalist		x	
Etsy.de	Generalist	x		
Eventim.de	Veranstaltungen	x		
Expedia.de	Reisen		x	
Facebook.de	Generalist		x	
Filmstarts.de	Filme	x		
Foodspring.de	Supplements	x		
de.Gearbest.com	Generalist			
Getyourguide.de	Veranstaltungen	x		
Google Play Store	Apps			x
Google.de	Generalist		x	
Google.de/shopping	Generalist	x	x	
Groupon.de	Generalist		x	
Holidaycheck.de	Reisen		x	
Hornbach.de	Baummarktartikel	x		
Idealo.de	Generalist	x	x	
Imdb.com	Filme			
Jameda.de	Ärzte		x	
Kennstdueneinen.de	Handwerker		x	
Kino.de	Filme	x		
Kleiderkreisel.de	Kleidung		x	
Lieferando.de	Restaurants		x	
Maxdome.de	Filme	x		

Portal	Branche (Schwerpunkt)	Antworten erhalten zu den Fragebögen		
		Produkt- bewertungen	Unternehmens- bewertungen	App- Bewertungen
Mediamarkt.de	Elektronikartikel	x		
Medizinfuchs.de	Medikamente	x	x	
Mindfactory.de	Elektronikartikel	x		
Mobile.de	Autohändler		x	
Moviepilot.de	Filme	x		
Myhammer.de	Handwerker		x	
Myprotein.de	Supplements			
Notebooksbilliger.de	Elektronikartikel	x		
Obi.de	Baumarktartikel	x		
Opodo.de	Reisen		x	
Otto.de	Generalist	x		
Pearl.de	Elektronikartikel	x		
Quandoo.de	Restaurants		x	
Rebuy.de	Elektronikartikel	x		
Saturn.de	Elektronikartikel	x		
de.Shein.com	Kleidung			
Shopapotheke.de	Medikamente	x		
Steampowered.com	Spiele			x
Tripadvisor.de	Reisen		x	
Trustpilot.de	Generalist		x	
Werkenntdenbesten.de	Anwälte/Finanzen		x	
Whofinance.de	Anwälte/Finanzen		x	
Yelp.de	Generalist		x	
Zalando.de	Kleidung	x		
Summe der Antworten:		28	36	3

Befragte Bewertungsvermittler

Betreibendes Unternehmen	Website(s)	Postanschrift in
AGRON Consult LTD & Co. KG	echte-bewertungen-kaufen.eu	Österreich
Alsayegh Media LLC	deutschebewertungen.com	VAE
ATTAG GmbH	attag.de	Deutschland
BD Management Ltd.	bewertungsdocus.com	Malta
Bewertungs-King	bewertungs-king.de	Deutschland
Bewertungs-Support	bewertungs-support.com	Deutschland
Brand Pioneers UG & Co. KG	market-pioneers.de	Deutschland
Christian Vondru Digital Marketing	bewertungenkaufen.at	Österreich
Cobus Company GmbH	test.a-z-rezensionen.de	Deutschland
Content-Werkstatt	content-werkstatt.com	Deutschland
Feedbackers UG	myfeedbackshop.de	Deutschland
Fivestar Services UG	fivestar-marketing.net	Deutschland
FS Management GMBH	finostars.com	Österreich
getViews Consulting UG	apps-planet.de	Deutschland
Internet Marketing Mallorca S.L.	positive-bewertungen-kaufen.de, bewertungenkaufen.net, bewertungenkaufen.com, bewertungen-kaufen24.com	Spanien
Likergeiz	likergeiz.de	Deutschland
Lutendo LLC	lutendo.com	Georgien
NcomTel LLC	googlebewertungen.com	USA / Österreich
NECMAI	necmai.com	Deutschland
O.M. INT. Fivestar Marketing LTD	star-builder.com	Zypern
PowerStar Ltd.	testerjob.net	Malta
Profiausbau Aachen UG	media-aachen.de	Deutschland
Review Intelligence GmbH	amarate.me	Deutschland
SatiSupp International LTD	vivarate.com	Zypern
ShopDoc GmbH	shopdoc.de	Deutschland
Social Media Market	socialmediamarket.de	Deutschland
SRM Global Ltd.	amztigers.de, goldstar-marketing.net	Malta
Vorwärts GmbH	clubderprodukttester.de	Deutschland
W3 Internet Marketing Lorenz GmbH	seo-hilfe24.de, suchmaschinenoptimierung- muenchen.com	Deutschland
WeComBlue Ltd.	wecomblue.com	Hongkong