

Das Prinzip Metacritic und seine Folgen

HILFREICH UND STARK UMSTRITTEN

Nur wenige Zahlen werden in der Spielebranche so heiß diskutiert und sind so einflussreich wie der Score auf Metacritic. Daran ändern selbst die fundamentalen Probleme der Plattform nichts. Von Martin Dietrich



Als der Videospieldesigner Adam Sessler 2009 die Bühne auf der Entwicklerversammlung GDC betritt, ist er sichtlich geladen. Sessler hält einen Kurzvortrag mit dem Titel »Fuck Metacritic!« Kürzlich habe er einen befreundeten Entwickler getroffen, der sich über einen Test bei Sesslers Fernsehsendung Xplay furchtbar aufgeregt habe. Weniger die zwei von fünf Punkten, die sein Tester vergab, seien das Problem gewesen, erzählt Sessler. Stattdessen habe die Bewertungsplattform

Metacritic ihre Einschätzung in die hauseigene 100er-Skala übersetzt, wodurch aus zwei Punkten nun 40 wurden. Der Gesamt-Score sei in der Folge ordentlich nach unten gerutscht, was finanzielle Konsequenzen nach sich ziehen könne. »Es ist kein Geheimnis, dass Metacritic als Werkzeug dient, um Vergütungen vorzuenthalten. Und das ist schlichtweg Blödsinn«, ruft Sessler unter dem Applaus des Publikums. Zudem sollen sich Publisher dafür schämen, führt der Xplay-Chefredakteur fort, dass sie Metacritic dazu nutzen würden, um die Qualität eines Spiels zu bestimmen. »Ihr verlasst euch auf uns? Wirklich? Wir könnten gemeinsam wahrscheinlich nicht mal ein Zelt

aufbauen und sollen nun mehr oder weniger einstimmig über die Qualität eurer Arbeit urteilen?« Sesslers emotionaler Ausbruch ist bei Weitem kein Einzelfall. 2011 beklagt Glen Schofield, damaliger Chefentwickler bei Sledgehammer Games, den niedrigen User-Score bei Modern Warfare 3 und wirbt öffentlich dafür, dass Fans doch mehr positive Bewertungen abgeben sollen.

Im selben Jahr streicht Metacritic nach Protesten den fragwürdigen Karriere-Score. Der Wert gab einen Durchschnitt an – zusammengesetzt aus allen Spielen, an denen ein Entwickler oder eine Entwicklerin beteiligt war. Und 2021 beschwert sich John Garvin, Creative Director von Days Gone, über die Zahlenversessenheit seines Publishers. »Die Realität bei Sony lautet, dass Metacritic alles ist«, sagt er. Ein Metascore von 70 bedeutet demnach eine kurze Laufbahn als Creative Director, so Garvin.

Bei einer Plattform, die es sich zur Aufgabe gemacht hat, die wichtigsten Pressebewertungen zu Filmen, Serien und Videospiele zu sammeln und daraus eine scheinbar objektive Zahl zu errechnen, wirken diese Meinungen dazu zunächst harsch. Angesichts der enormen Bedeutung, die Metacritic und ähnliche Konkurrenzplattformen immer noch besitzen, ergibt sich jedoch ein anderes Bild. Metacritic aggregiert nicht nur stumpf die Zahl am Ende einer ausführlichen Kritik. Die Webseite prägt überdies den Diskurs über Videospiele und kann manchmal den Entwicklungsprozess selbst beeinflussen. Wie genau, das haben wir für diesen Report recherchiert. Denn hinter den



User-Scores bei Metacritic zeichnen sich oft durch eine hohe Emotionalität aus. Trotz hoher Qualität enttäuschte Diablo 4 viele Fans, weil es ihrer Meinung nach aktuell an der Langzeitmotivation und abwechslungsreichen Seasons fehlt. Der User-Score liegt bei 2,2 und damit tief im roten Bereich.

Kulissen geht bei Metacritic nicht immer alles mit rechten Dingen zu.

Die Mission von Metacritic

Metacritic wird 2001 von den Jura-Absolventen Marc Doyle und Jason Dietz sowie Doyles Schwester Julie Roberts gegründet. Wie Doyle 2010 in einem Interview mit der San Diego Union-Tribune erklärt, waren sie vom Internet und dessen Fähigkeit, viele verschiedene Meinungen zu vereinen, fasziniert. Statt einzelner Expertenmeinungen soll die gebündelte Fusion von Fachwissen als Verkaufsberater dienen. »Die Mission von Metacritic besteht darin, die Meinungen der besten Kritiker zusammenzubringen, damit die einzigartige und persönliche Meinung eines einzelnen Kritikers die Kaufentscheidungen nicht übermäßig beeinflussen kann«, sagt Doyle. Insbesondere freue es ihn, dass die Branche neben den reinen Verkaufszahlen nun auf eine zweite Metrik schaue, um die Qualität ihrer Produkte zu ermitteln. Und das scheint die Industrie in der Tat schnell umzusetzen: 2004 implementiert das Filmstudio Warner Bros. eine zusätzliche Klausel in seine Lizenzverträge. Publisher, die basierend auf einem Warner-Film ein Spiel herausbringen wollen, müssen Mindestwerte bei Metacritic und anderen ähnlichen Seiten wie Game Rankings erreichen. Landen sie darunter, muss der Publisher höhere Lizenzkosten leisten. Ungefähr im selben Zeitraum beginnen Activision und Take-Two Bonuszahlungen an Mitarbeitende daran festzumachen, ob Titel einen vorher festgelegten Score erreichen. Begründet wird dies mit der angeblichen Strahlkraft, mit der Metacritic und andere Aggregatoren auf die Absatzzahlen scheinen. Während der DICE-Awards 2007 erklärt



Gibt es bis heute keine Fortsetzung des Biker-Abenteuers Days Gone, weil Publisher Sony zu stark auf Metacritic-Bewertungen schießt? Das behauptete zumindest einer der Entwicklungschefs. Demnach sollen es alle Spiele schwer haben, die unter die magische Grenze von 85 Punkten fallen.

Activisions Marketingchefin Robin Kaminsky, dass sich für alle fünf Punkte über 80 bei Game Rankings die Verkäufe grob verdoppeln würden. Allerdings, so Kaminsky, sei der durchschnittliche Kritiker-Score allein nicht für einen Millionenerfolg verantwortlich. Vielmehr sei die magische Zahl erst im intelligenten Verbund mit effektivem Marketing und einer zielgruppengerechten Verkaufsstrategie wirklich erfolgreich.

Studien scheinen dies zunächst zu bestätigen. Eine Untersuchung der Full Sail University aus Florida von 2013 spricht von einer »engen Kopplung zwischen Umsatz und Metacritic-Score« über alle Genres und Plattformen hinweg. Ein definitiver Beweis wurde aber bis heute nicht erbracht. Dass GTA 5, Zelda oder Uncharted nicht nur bei vielen Kritikern gut ankommen, sondern auch Verkaufsrekorde aufstellen, muss nichts mit Metacritic zu tun haben. Die am besten bewerteten Titel auf solchen Plattformen sind häufig Spiele mit einer etablierten Fan-Ge-

meinschaft und einer dementsprechend großen Marketingmacht hinter sich. Zudem dürften die allermeisten professionellen Kritiker und Kritikerinnen einen sehr guten Riecher haben, welche Spiele eine möglichst breite Gruppe ansprechen. Anders gesagt: Metacritic listet zwar keine Verkaufszahlen, aber seine Statistik zeigt am Ende doch oft, welche Titel beliebt und damit am Markt außerordentlich performen können.

»Diesem Wert kannst du nicht vertrauen«

Trotz der fehlenden Fähigkeit als Business-Hellseher kann es sich daher für Spielestudios lohnen, Metacritic als Gradmesser im Blick zu behalten. »Wenn die ersten Reviews eintrudeln, dann guckt man sich die Wertungen dort schon an«, sagt Michael Hoss, Head of Product bei Deck13 Spotlight, der Publishing-Sparte der Frankfurter Firma. »Wir wollen schon bei einem gewissen Wert bleiben. Was das Publishing angeht, sind



Die Handkantensimulation God Hand von Resident-Evil-Erfinder Shinji Mikami kam zur Veröffentlichung bei der Presse mit einem Metacritic-Score von 73 eher mäßig an. Bei Spielern und Spielerinnen ist God Hand aber mittlerweile so etwas wie ein Kult-Hit. Der User-Score liegt bei satten 9,2 Punkten.



Chained Echoes ist eine Liebeserklärung an JRPGs der alten Schule und wurde 2023 als bestes Spiel beim Deutschen Computerspielpreis ausgezeichnet. Nach der Veröffentlichung war es aber auch kurzzeitig Opfer von Review Bombing auf Metacritic. Nach Beschwerden des Publishers schritt Metacritic ein.

wir bei allem, was 70 oder mehr Punkte erreicht, erstmal glücklich. Das ist auch gar nicht so einfach«, meint Hoss. Wenn sie von einem Entwicklerstudio eine Idee präsentiert bekommen, dann könne der Score älterer Spiele schon das Zünglein an der Waage sein, ob ein Publishing-Deal unterschrieben wird. Da aber Spotlight wie bei der Städtebausimulation High Rise City oder dem 2D-Rollenspiel Drova häufig mit Entwicklern zusammenarbeitet, die zum ersten Mal an einem Spiel werkeln, sei am Ende auch die Qualität der Idee entscheidend, erklärt Hoss weiter. Bonuszahlungen anhand vorher festgelegter Metascores gebe es bei ihnen nicht. Auch der Rest der Branche scheint nach jahrelanger Kritik davon Abstand zu nehmen. Nach dem schwierigen Release von Cyberpunk 2077 änderte beispielsweise CD Projekt sein Bonussystem.

Zudem hält Hoss von Deck 13 Spotlight die Art und Weise, wie Metacritic die Scores berechnet, für wenig seriös. Denn Metacritic wählt eigenständig aus, welche Publikatio-

nen in den Score einfließen. Darüber hinaus wird die Bewertung in eine 100er-Skala überführt und so vereinheitlicht, was schon Adam Sessler in seiner Wutrede kritisierte. Zu guter Letzt werden die einzelnen Bewertungen dann je nach Bedeutung der Publikation gewichtet. Wie genau die Formel dazu aussieht, darüber will Metacritic bis heute nicht offen sprechen. »Das macht den Durchschnittswert komplett sinnlos. Diesem Wert kannst du nicht vertrauen«, sagt Hoss, der das Vorgehen des Konkurrenten OpenCritic bevorzugt, der keinerlei Gewichtung vornimmt. Schaut man sich die zehn am besten bewerteten Spiele 2023 bei Metacritic und OpenCritic an, werden die Unterschiede besonders deutlich: Während mit Baldur's Gate 3, Tears of the Kingdom und Metroid Prime Remastered bei beiden Plattformen dieselben Spiele die Top 3 bilden, wenn auch in einer anderen Reihenfolge, wird es bei den darauffolgenden Plätzen undurchsichtiger. Metacritic listet aktuell beispielsweise Alan Wake 2, Sea of Stars und

Diablo 4 als Kritikerlieblinge, die bei OpenCritic wiederum komplett fehlen. Dafür hat der Konkurrent mit Videoverse und Turbo Overkill zwei eher unbekannte Indieschmuckstücke in seiner Bestenliste.

Bei User-Scores herrscht große Emotionalität

Doch nicht nur der Kritiker-Score hat es schwer, ein halbwegs nachvollziehbares Meinungsbild abzubilden. Auch gewöhnliche Nutzer und Nutzerinnen können bei Metacritic im Rahmen einer 10er-Skala ihre Gedanken zu einem Spiel beziffern. Der User-Score verhält sich dabei oft diametral zur Presse und folgt seiner ganz eigenen Logik. Zu diesem Ergebnis kam eine groß angelegte Studie der Unis in Aachen und Graz in Österreich, die 2019 alle damals verfügbaren schriftlichen Spiele-Reviews auf Metacritic statistisch analysierte. Über eine Million Bewertungen von Nutzern und professionellen Testern wurden in der Studie ausgewertet und miteinander verglichen. »Die Nutzer tendieren dazu, entweder ganz positiv oder ganz negativ zu schreiben und zu bewerten«, sagt Denis Helic, Professor für Data Science an der Uni Wien und Mitautor der besagten Studie. Laut der Untersuchung greifen in der ersten Woche nach Veröffentlichung die Nutzer am häufigsten zu einer Null oder Zehn. Erst nach etwa vier Wochen normalisiere sich das Bewertungsschema. »Meine Vermutung wäre, dass die Spieler und Spielerinnen es deutlich emotionsgeladener betrachten«, sagt Helic. Die professionellen Kritiker würden dahingehend viel häufiger die ganze Breite der Bewertungsskala nutzen und in einem distanzierten Stil über ein Spiel schreiben.

Wenn sich innerhalb kürzester Zeit vor allem negative User-Bewertungen ansam-



Eine große Debatte über Metacritic entbrannte, nachdem bekannt wurde, dass Obsidian Entertainment für Fallout: New Vegas eine Bonuszahlung aufgrund eines einzigen Metacritic-Punktes verwehrt wurde. Heute gilt das Rollenspiel als einer der besten Vertreter des Genres.



Jede Bewertungsplattform ermittelt den Durchschnitt aller Pressekritiken unterschiedlich. In der Liste der besten Spiele 2023 taucht der extrem brutale und extrem schnelle Shooter Turbo Overkill bei Metacritic nicht auf, wohingegen er bei OpenCritic Platz zehn belegt.

lerinnen, die den DS-Titel womöglich aus ihrer Kindheit kennen, sind da deutlich nachgiebiger und vergeben für den Titel eine Traumwertung von 9,1 Punkten.

Wie sinnvoll ist Metacritic heutzutage noch?

All diese Beispiele zeigen, dass Metacritic trotz der Zielvorgabe des Co-Gründers Marc Doyle nicht einfach nur die Meinungen der besten Kritiker zusammengebracht hat. Angesichts stetig wachsender Service-Spiele, die alle paar Wochen oder Monate Features entfernen und hinzufügen, scheint eine absolute, in Stein gemeißelte Note mit Anspruch auf Vollständigkeit obsolet.

Da Metacritic nicht erlaubt, eine abgegebene Bewertung zu revidieren beziehungsweise anzupassen, bleibt der Kritiker-Score eine Momentaufnahme, eingefroren kurz nach der Veröffentlichung. Oder hat sich jemand schon mal über den niedrigen Metascore von Welthits wie League of Legends oder Fortnite beschwert? Auch viele Mobile-Spiele werden von Metacritic und großen Teilen der Presse ignoriert. Einer finnischen Studie zufolge hat gerade mal ein Viertel der 1.000 erfolgreichsten iPad-Spiele in den USA überhaupt einen Metascore.

Die Zeiten, in denen Metacritic die Verkäufe eines Spiels ankurbeln kann, sind laut Matthew Karch, Chef von Saber Interactive, sowieso vorbei. Trotz mäßiger Kritiken habe sich Evil Dead: The Game immer noch richtig gut verkauft. Karch nennt im Rahmen eines Finanzreports der Embracer Group stattdessen Social Media und Influencer als weitaus größere Einflussfaktoren. Gerade Multiplayer-Spiele, die von viralen Momenten leben, können wohl getrost auf eine 95 bei Metacritic verzichten. Ob die Plattform selbst allerdings irgendwann verschwinden wird, das dürfte weniger sicher sein. Denn solange eine Zahl am Ende eines langen Textes oder eines Videos als Qualitätssiegel gilt, werden wir wohl versuchen, einen allgemeinen Durchschnitt daraus zu errechnen. Dafür liebt der Mensch das Vergleichen, Benoten und natürlich auch das Streiten darüber anscheinend dann doch zu sehr. ★

melnen, spricht man vom sogenannten Review Bombing. Der Sturm der Entrüstung kann einerseits wüten, weil Titel wie Battlefield 2042 oder Diablo 4 (User-Score jeweils lediglich 2,2 Punkte) Fans inhaltlich enttäuschen. Auch werden Metacritic und Co. immer öfter genutzt, um die generelle Entwicklung einer Marke oder aggressive Monetarisierungsstrategien wie bei NBA 2K24 (User-Score 2,0 Punkte) abzustrafen.

Andererseits sind Videospiele seit längerem Teil rechter Kulturkämpfe. Wenn es Spiele wie Horizon Forbidden West: Burning Shores (User-Score 4,5 Punkte) oder The Last of Us 2 (User-Score 5,8 Punkte) wagen, abseits etablierter Schablonen ihre Geschichten diverser zu erzählen, dann folgen oft auch koordinierte Hasskampagnen gegen die Entwicklerstudios. Review Bombing ist 2020 ein so großes Problem auf der Plattform, dass sich Metacritic dazu genötigt sieht, eine Zeitsperre einzuführen. Seitdem dürfen User erst 36 Stunden nach Veröffentlichung eines Titels eine Bewertung abgeben. Komplett verschwunden ist das Phänomen dadurch aber nicht.

Review-Bombing kann auch positive Effekte haben

Auch das von Deck13 betreute Oldschool-Rollenspiel Chained Echoes, das 2023 als bestes deutsches Spiel beim DCP ausgezeichnet wird, ruft eines Tages mies gelaunte Internetnutzer auf den Plan. Der User-Score für die PlayStation- und Switch-Fassung sackt in der Folge auf unter sechs Punkte ab. Den Grund für den Zorn kennt selbst Michael Hoss bis heute nicht, da einfach nur massenhaft schlechte Bewertungen ohne Wortmeldungen eintrudelten. Er vermutet jedoch, dass es an einer fehlenden Sprachfassung für den lateinamerikanischen Markt gelegen haben könnte. Statt sich aber mit dem Schönheitsfehler abzufinden, beschwert sich Hoss auf X (vormals Twitter) öffentlich über die unerklärliche Schwemme an Negativität. »Ehrlicherweise war das kalkuliert, was wir gemacht haben«, gibt Hoss zu. Die Hommage an JRPG-Klassiker hat zu diesem Zeitpunkt für ein Indie-Spiel schon reichlich Wohlwollen gesammelt

und sich bei Fans einen Namen gemacht. Trotzdem bringt der Post nochmal viel Aufmerksamkeit. Größere amerikanische Webseiten schreiben über den Fall, auf X wird über den fehlenden Schutz vor Review Bombing diskutiert. In den nächsten Tagen steigt das Interesse an Chained Echoes deutlich an, einige tausend verkaufte Kopien mehr habe die »kalkulierte Empörung« eingebracht, schätzt Hoss. »Wenn die Nutzer das System ad absurdum führen, dann ist es doch mein gutes Recht als Hersteller dieses System ebenfalls ad absurdum zu führen.« Kurze Zeit nach der Aufregung reagiert Metacritic und entfernt die Griesgrame.

Neben größerer Emotionalität spielen auch Nostalgie auf Metacritic eine große Rolle, zeigt die Studie von Denis Helic und seinen Kollegen aus Österreich. Während die klassische Presse maximal neun Tage nach Release ihre Wertung abgibt, dauert es bei normalen Nutzern deutlich länger. Die Hälfte aller User-Reviews erscheint mehr als zwei Monate nach Veröffentlichung. Und wohingegen die anfängliche Resonanz bei Nutzern häufig eher schlecht ist, wird sie mit fortschreitender Zeit immer besser. Sobald ein Spiel älter als fünf Jahre ist, nehmen die positiven Bewertungen sogar deutlich zu. So hat es Pokémon Mystery Dungeon: Explorers of Sky mit einem Kritiker-Score von 54 sehr schwer bei der Presse. Die Spieler und Spie-



Review Bombing muss nicht immer damit zusammenhängen, dass ein Spiel technisch oder spielerisch enttäuscht. Bei Horizon Forbidden West: Burning Shores reichte bereits die Implementierung einer homosexuellen Beziehung aus.