

Serie Evaluation (2): Der getestete Test – Welche Kriterien sind wichtig, um aussagekräftige Ergebnisse zu erzielen?

Stellen Sie sich vor, Sie besuchen zum Beispiel eine große Ausstellung. Nun hält Sie während Ihres Aufenthalts ein uniformierter und bewaffneter Polizist an, der Sie auffordert, ihm für ein Interview Rede und Antwort zu stehen. Eigentlich haben Sie ja etwas anderes geplant, als Ihre Zeit mit Frage-Antwort-Spielchen zu verbringen, daher fragen Sie nach, ob das denn unbedingt sein müsste. Als Antwort erhalten Sie ein leicht schneidendes, klares: „Was glauben Sie denn, warum ich hier bin?“

Diese Geschichte ist nicht erfunden, sondern basiert auf einer wirklichen Begebenheit (Moscardo 2006: S. 26). Im Ergebnis ließen sich alle der auf diese Art Angesprochenen interviewen, vermutlich beeinflusst durch die (angenommene) Respektsperson. Die Vermutung liegt deshalb nahe, weil bei den Kollegen, die ebenfalls Interviews in besagter Ausstellung durchführten, im Durchschnitt 25% der Angesprochenen das Interview verweigerten – im Gegensatz zum „Polizisten“, der keine einzige Verweigerung vorzuweisen hatte. Was war geschehen? Der „Polizist“ war zwar tatsächlich einer, allerdings war die Interviewertätigkeit lediglich ein Zubrot für ihn. Aufgebrezelt in voller Uniform in Verbindung mit autoritärem Auftreten „gewann“ er alle Be-

fragen für das Interview – und verfälschte dabei möglicherweise die Befragungsergebnisse. Die Wahrscheinlichkeit, dass Gäste „sozial erwünscht“ geantwortet haben, liegt bei einem derartigen Vorgehen nahe. „Sozial erwünscht“ bedeutet, dass interviewte Gäste das antworten, was sie annehmen, dass es der Interviewer hören will und nicht unbedingt das, was sie wirklich meinen. Dies ist bei Befragungen ein weit verbreitetes Problem, das aus vielen Gründen zum Tragen kommen kann. Daher sollten Interviewer grundsätzlich so neutral wie möglich auftreten, um zu vermeiden, die Befragten bereits durch die Frageformulierung zu beeinflussen (Kleidung, Ansprache, Auftreten, etc.). Während die sogenannte „soziale Erwünschtheit“ eine von mehreren Fragen

der Gültigkeit in der Besucherforschung ist, gibt es auch noch Aspekte, die sich mit der Verlässlichkeit und der Repräsentativität einer Erhebung befassen. Diese drei Punkte: Gültigkeit, Verlässlichkeit und Repräsentativität sind die zentralen Kriterien, an denen fest gemacht werden kann, ob ein Test ordentlich konzipiert und durchgeführt wurde – oder nicht. Um diese drei Kriterien, gleichermaßen der „Eich-test“ für Evaluationen, geht es hier.

Gültigkeit

Das für eine Evaluation ausgewählte Erhebungsinstrument sollte sinnvoll sein. Wenn das nicht der Fall ist, ist die Evaluation im hier beschriebenen Sinne nicht „gültig“. Zentral ist also eine sinnvolle Methoden-

**Composit- & Metallbau
Kullmann GbR**
Warteberg 12 - 37327 Leinefelde
Tel. 03605 501720 Fax 03605 512590
e-mail: kullmannmetallbau@t-online.de

Kaninchengehege Sababurg

Alles aus einer Hand:
Lieferung und Montage von Zaunanlagen, kompletten Gehegen, E-Zaunanlagen, Vergitterungen, Übernetzungen;
Lieferung von Gehegen und Voljieren im Baukastensystem zur Eigenmontage - auf Wunsch mit Konstruktion nach Ihren Vorgaben;
Lieferung von einzelnen Komponenten wie Schieber jeder Art, Pylone - auch teleskopierbar, Fütterungsanlagen, Trennwände

DPZ Göttingen

wahl. Ein überspitztes(!) Beispiel wäre, kleinen Kindern bei Geburtstagsfeiern im Tierpark Fragebögen und Stift zwecks schriftlicher Rückmeldung auszuhändigen. Der Rücklauf, d. h. ausgefüllte Bögen, wären nicht wirklich zu erwarten.

Auch macht es keinerlei Sinn, zu versuchen, mittels Beobachtung aus 15 Metern Entfernung, Gespräche von Gästen mitzubekommen, mit dem Ziel, herauszubekommen, wie groß der Gesprächsanteil ist, der sich mit Tieren und der Einrichtung im engeren Sinne beschäftigt. Gleichzeitig kann es, je nach Untersuchungsziel im-

mens wichtig sein, Gesprächsinhalte mitzubekommen (siehe Kasten 1).

Offen ausgelegte Besucherbücher, in die jedermann eintragen kann, was er möchte, sind ebenfalls mit Vorsicht zu interpretieren (Terolog 2006: S. 38f.). Vor allem sollten normalerweise keine abschließenden Entscheidungen auf dieser Basis getroffen werden. Ein mir bekannter Zoo rechtfertigt beispielsweise eine intensivere Beschilderung damit, dass Gästebefragungen dies ergeben hätten. Auf intensive Nachfrage stellte sich heraus, dass dieses Ergebnis auf Gästebuchauswertungen zurückzuführen

war. Aussagen in Gästebüchern (oder ausliegenden Fragebögen) können durchaus ein Grund sein, die Gäste einmal konkret in einer bestimmten Richtung zu befragen. Spezifische Maßnahmen jedoch allein auf Grund sporadischer Gästebucheinträge einer in Größe und Eigenschaften unbekanntenen Gruppe von Gästen einzuleiten, ist zumindest als schwierig zu beurteilen.

Kurz: Es gibt verschiedenste Erhebungsmethoden, schriftliche Befragungen sind nur eine davon. Die Wahl der geeigneten Methode hängt vom Untersuchungsgegenstand ab. Wichtig ist außerdem ein möglichst neutrales Befragungsumfeld zu schaffen, um soziale Erwünschtheit zu vermeiden.

„Untersuchung zeigt, dass Zoobesucher wenig Zeit mit der Betrachtung von Tieren verbringen – sogar bei fast ausgestorbenen Arten

(veröffentlicht am 20.04.2006, Kansas State Collegian)

Zoos werben aufdringlich mit ihrem Bildungsbestreben, jedoch sind sie für die Personen, die den Zoo besuchen um "mal aus dem Haus zu kommen" lediglich eine leichte Ablenkung. Dale Marcellini, ein Verwalter des nationalen Zoos in Washington D.C., führte eine Studie über Zoobesucher durch, in welcher er mit mehreren Kollegen mehr als 700 Menschen beobachtete, verfolgte und ihnen, über ein paar Sommer hinweg, zuhörte.

Sein Fazit war, dass Zoos kaum mehr als die Hintergrundkulisse für die Beschäftigung mit persönlichen Themen sind. Bei den Gesprächen der Besucher ging es nicht um die Tiere, sondern um ihr eigenes Leben. Wenn die Menschen tatsächlich Bemerkungen zu einem Tier machten, waren es meist abfällige Bemerkungen. Die Studie zeigte, dass 60% der verbrachten Zeit damit verbracht wurde, rumzulaufen, fast 10% der Zeit wurden mit Essen verbracht, weitere Zeitanteile wurden verwendet, um auszuruhen, auf Toilette zu gehen oder einzukaufen.

Die Besucher verbrachten weniger als eine Sekunde pro Schlange und eine Minute bei den Löwen. Der Davidshirsch, welcher als ausgestorben gilt, wenn das letzte der noch in Gefangenschaft lebenden Tiere stirbt, wurde lediglich 27 Sekunden beachtet.

Es sind nicht nur die Besucher, die desinteressiert sind. So sagte zum Beispiel sogar ein ehemaliger Direktor des berühmten Zoos in Atlanta über die Tiere: „Die sind das Letzte worüber ich mir Sorgen mache, bei all den anderen Problemen.““ (frei übersetzt)

Kasten 1: Beispiel für eine Untersuchung unter anderem der Gesprächsinhalte von Zoo-Gästen (<http://bigcatnews.blogspot.de/2006/04/study-finds-zoo-visitors-spend-little.html> (Stand: 29.08.2013))

Wer sind die Jugendlichen? Jugend zwischen 10 Jahren und „forever young“

- 12 - 28 Jahren (Jugendstiftung)
- 14 - 18 Jahren (Porwol)
- 14 - 21 Jahren (Danckwörtt)
- 14 - 25 Jahren (Wettbewerb „Jugend reist und lernt Europa kennen“)
- 14 - 27 Jahren (Kinder- und Jugendhilfegesetz)
- 14 - 29 Jahren (RA StfT/Marktanalyse Grund & Jahr; EMNID)
- 15 - 24 Jahren (European Trave Monitor)
- „forever young“ ? (Medien)

[Quelle: Orlovitz-Wessely, Anita (1997): Viel gereist und nichts gelernt? Wirkungen von Auslandsreisen bei Jugendlichen. Thomas Mord Akademie (Hg.), Benzberger Studien 9.]

Repräsentativität

Repräsentativ ist eine Erhebung dann, wenn die befragten, beobachteten oder anders untersuchten Gäste hinsichtlich ihrer Zusammensetzung dem Normalpublikum der Einrichtung entsprechen.

Das bedeutet erstens, dass prozentual verteilt möglichst dieselbe Menge der verschiedenen Altersgruppen vertreten sein sollte, wie dies auch im Alltag des Parks der Fall ist. Dazu müssen Altersgruppen festgelegt werden, die natürlich nicht von Erhebung zu Erhebung zu ändern sind, was die Vergleichbarkeit der Daten unmöglich machen würde (siehe Folgende: Verlässlichkeit). Dabei sollten Sie sich nicht von verschiedenen Altersgruppendefinitionen verwirren lassen (Abb. 1). Wichtig ist, dass Ihr Wildpark weiß, was er für seine Zwecke unter Zielgruppen wie z. B. „Kinder“, „Jugendliche“, „junger Familien“, „Senioren“, etc. versteht.

Zweitens sollte die Erhebung möglichst alle Öffnungszeiten abdecken (Abb. 2). Dies gilt vor allem, wenn zu vermuten oder bekannt ist, dass bestimmte Gästegruppen gerne zu bestimmten Tages- oder auch Jahreszeiten kommen. Überspitztes(!) Beispiel: Wenn Sie wissen, dass Senioren (bspw. mit einem Alter ab 70 definiert) gerne vormittags zu Besuch kommen, verzerrt es das Ergebnis hinsichtlich der Repräsentativität, wenn aus Gründen der Arbeitsorganisation nur Nachmittags befragt werden kann.

Und drittens sollten genügend Gäste untersucht werden, um zu gewährleisten, dass die Daten verallgemeinerbar sind. Wenn eine Erhebung auf Verallgemeiner-

Abb. 1 - Wer ist jugendlich

Liste von möglichen Beobachtungsperioden

Montag 3. Juni	Dienstag 4. Juni	Mittwoch 5. Juni	Donnerstag 6. Juni	Freitag 7. Juni	Samstag 8. Juni	Sonntag 9. Juni
10:00 - 11:00	10:00 - 11:00	10:00 - 11:00	10:00 - 11:00	10:00 - 11:00	10:00 - 11:00	10:00 - 11:00
11:00 - 12:00	11:00 - 12:00	11:00 - 12:00	11:00 - 12:00	11:00 - 12:00	11:00 - 12:00	11:00 - 12:00
12:00 - 13:00	12:00 - 13:00	12:00 - 13:00	12:00 - 13:00	12:00 - 13:00	12:00 - 13:00	12:00 - 13:00
13:00 - 14:00	13:00 - 14:00	13:00 - 14:00	13:00 - 14:00	13:00 - 14:00	13:00 - 14:00	13:00 - 14:00
14:00 - 15:00	14:00 - 15:00	14:00 - 15:00	14:00 - 15:00	14:00 - 15:00	14:00 - 15:00	14:00 - 15:00
15:00 - 16:00	15:00 - 16:00	15:00 - 16:00	15:00 - 16:00	15:00 - 16:00	15:00 - 16:00	15:00 - 16:00
16:00 - 17:00	16:00 - 17:00	16:00 - 17:00	16:00 - 17:00	16:00 - 17:00	16:00 - 17:00	16:00 - 17:00
17:00 - 18:00	17:00 - 18:00	17:00 - 18:00	17:00 - 18:00	17:00 - 18:00	17:00 - 18:00	17:00 - 18:00

Beginnen Sie hier und wählen Sie dann jede siebte Stunde.

Abb. 2 - Staffelfung von Öffnungszeiten

barkeit aus ist, dann ist es sinnvoll, von mindestens 100-150 Gästen Daten zu erheben. Evaluationen, die für viele oder vielleicht sogar alle Wild- und Tierparks interessant sind, könnten in Abstimmung mit Kollegen in mehreren Parks im selben Zeitraum durchgeführt werden, da sich so die Repräsentativität deutlich verstärken lässt.

Verlässlichkeit

Ergebnisse sollten wiederholbar sein. Wenn mehrfache Erhebungen unter gleichen Bedingungen mit denselben Gästen jedes Mal zu anderen Ergebnissen führen, dann ist die Untersuchung nicht „verlässlich“. Verlässlichkeit ist eine Grundlage dafür, dass Evaluationen auch „gültig“ (s. o.) sind. Wenn zum Beispiel verschiedene Mitarbeiter an einer Erhebung mitarbeiten und dabei unterschiedlich vorgehen, dann ist es nicht verwunderlich, wenn auch die Ergebnisse unterschiedlich ausfallen. Grundsätzlich kann es viele Gründe für eine mangelnde Verlässlichkeit geben - fehlerhafte Durchführung, unzureichende Datenmenge, unkorrekte Auswertung, etc. Kurz: Dies bedeutet, dass Konzeption, Test, Durchführung und Auswertung einer Untersuchung unter vergleichbaren Bedingungen stattfinden müssen, um zu verlässlichen Ergebnissen zu kommen.

Fazit

Wie wichtig technisch sauber durchgeführte Evaluationen sind, zeigt sich immer wieder. Ein aktuelles, prominentes Beispiel dürfte die massive methodische Kritik (Marion et al. 2010) an einer in den USA „hochgejubelten“ Studie über positive Effekte bei Einstellungsänderungen von Zoobesuchern (Falk et al. 2007) sein. Zu den aufgezeigten Mängeln zählen die Kritiker neben unangemessenen Vergleichs-

gruppen und fehlenden Vorher-Nachher-Tests auch Experimentator-Effekte oder die Abfrage nicht nur von Meinungen und Einstellungen sondern auch tatsächlichem Wissen, etc.

Aufwändige Arbeiten sind häufig im Alltag tiergärtnerischer Einrichtungen (noch) nicht realisierbar. Wünschenswert wäre es dennoch, wenn der Erhebung von Daten im informellen und formellen Bildungsbereich künftig deutlich mehr Gewicht zukäme und sich Gästeuntersuchungen nicht fast ausschließlich auf Marketingthemen beziehen. Denn eine solide Datengrundlage ist ein hervorragender Baustein für eine entsprechend sinnvolle und attraktive Angebotsgestaltung, was wiederum attraktiv für Gäste einerseits und potenzielle Mittelgeber andererseits wirken kann.

Eine gültige, repräsentative und verlässliche Evaluation bietet dafür eine solide Grundlage.

Verfasser: Lars Wohlers, KON-TIKI

Literatur

Falk, J. H. et al.: Wy zoos & aquariums matter: Assessing the impact of a visit to a zoo or aquarium. Silver Spring, MD: Association of Zoos & Aquariums.

Marion, Lori et al.: Do Zoos and Aquariums Promote Attitude Change in Visitors? A Critical Evaluation of the American Zoo and Aquarium Study. In: Society and Animals Vol. 18 / 2010 / p. 126 – 138.

Moscardo, Gianna: Visitor Studies Bloggers: A Funny Thing Happened ... in: Visitor Studies Association (VSA, pub.): Visitor Studies Today Vol. 9 / Issue 3 / 2006 / p. 25-29.

O'Connor, Jennifer (o. J.): Study finds zoo visitors spend little time viewing animals, even nearly extinct species. In: http://ccaawps.ca/answers/details/doesnt_the_guzoo_have_educational_value (Stand: 28.08.2013)

Terolog (2006): Summative Evaluation der Greenpeace-Ausstellung „Zertifikat Nr. 000358/ verstrahlt – verdrängt – vergessen“ zum 20. Jahrestag der Tschernobyl-Katastrophe in Hamburg auf Kampnagel [K]. (unveröff., im Auftrag von Greenpeace Deutschland e. V.).

Wohlers, Lars (2012): Was interessiert Zoo- und Wildpark-Gäste? In: Deutscher Wildgehege-Verband (2012): WildTierZeit, Ausgabe 2/2012, S.32.



Maisel Plüsch

Naturgetreue und qualitativ hochwertige Plüschtiere aller Art

Wir beliefern Zoos, Tierparks und Vivarien in ganz Deutschland. Alle unsere Plüschtiere entsprechen der europäischen Sicherheitsnorm für Spielzeug DIN EN 71.



Wellstr. 1
59439 Holzwickede
Telefon: 02301/945 40 43
Telefax: 02301/945 40 47
Mail: info@maisel-pluesch.de
Web: www.maisel-pluesch.de

