

## V. Die Arbeit an der Evidenz

---

Ich habe im vorigen Kapitel die Sanftheit der Verhaltensbeeinflussung durch Nudges herausgearbeitet (vgl. Bröckling 2017). Umgebungen werden so verändert, dass sie in die Aktivitäten der Beeinflussten hineingezogen, und so zu Umwelten im Sinne Uexkülls werden können. Damit wirken Nudges der Tendenz nach präreflexiv und die Verhaltensbeeinflussung erfolgt so un bemerkt, dass sich das Konzept von Anfang an mit Manipulationsvorwürfen auseinandersetzen musste (vgl. Schmidt/Engelen 2020: 7). In diesem Kapitel untersuche ich, vor welche epistemischen Herausforderungen die Praktiker:innen durch die Sanftheit ihrer Regierungsbemühungen gestellt sind. Denn weil die Beeinflussung so sanft erfolgt, dass die Beeinflussten es nicht merken, müssen die Praktiker:innen irgendwie nachweisen, dass ihre Eingriffe in Umgebungen den gewünschten Effekt haben. Sie müssen Evidenz für die Wirksamkeit ihrer Verhaltensbeeinflussung beschaffen, ohne die der kleine Stupser eben nur ein kleiner Stupser bliebe. Diese Evidenz wird meist in Form eingängiger Balkendiagrammen dargestellt, denen die Wirksamkeit der jeweiligen Intervention leicht abgelesen werden kann.

Ich folge in diesem Kapitel einem Argument von Harold Garfinkel, demzufolge die Art, wie über ein Phänomen berichtet wird, nicht losgelöst vom Phänomen ist, sondern integraler Bestandteil desselben: »Reporting procedures, their results, and the uses of these results« – so schreibt er – »are integral features of the same social orders they describe.« (Garfinkel 1967: 192) Demzufolge lässt sich aus der Art, wie die Praktiker:innen über Nudging berichten, einiges über Nudging selbst lernen. In einem Verständnis des Nudges als formativem Objekt ist die Evidenz damit nicht losgelöst von Problem und Lösung zu verstehen. Sie gehört zu einem gelungenen Nudge genauso, wie die beiden anderen Teilobjekten. Ohne sie wäre die große Wirkung die dem kleinen Stupser zugesprochen wird, nicht »detectable, countable, recordable, reportable, tell-a-story-about-able, analyzable – in short *accountable*« (ebd.: 33).

Und ohne die Evidenz wäre es nicht möglich, Geschichten eines gelungenen Nudges wie die folgende zu erzählen, auf die ich bereits im vorigen Kapitel eingegangen bin, um Uexkülls Umweltbegriff zu illustrieren. Nun widme ich mich ihr erneut, weil sich von ihr auch über die epistemischen Herausforderungen des Nudging viel lernen lässt:

One of the most scenic urban highways in the world is Chicago's Lake Shore Drive, which hugs the Lake Michigan coastline that is the city's eastern boundary. The drive offers stunning views of Chicago's magnificent skyline. There is one stretch of this road that puts the drivers through a series of S curves. These curves are dangerous. Many drivers fail to take heed of the reduced speed limit (25 mph) and wipe out. In September 2006, the city adopted a new strategy for slowing traffic. It painted a series of white lines perpendicular to the traveling cars. The lines progressively narrow as drivers approach the sharpest point of the curve, giving them the illusion of speeding up, and nudging them to tap their brakes. (Thaler/Sunstein/Balz 2013: 433)

Diese Geschichte ist ein typisches Beispiel dafür, wie Nudging als neue, innovative und außerordentlich wirksame Form der Verhaltensbeeinflussung präsentiert wird. Durch ungewöhnliche und vergleichsweise kleine Veränderungen der Umgebung werden große Effekte erzielt. Nur ein paar auf die Straße gemalte weiße Linien werden so in die Umwelt der abgelenkten Fahrer:innen hineingezogen, dass diese quasi automatisch auf die Bremse treten. Mehr noch: Nudging beweist dort seine Wirksamkeit, wo andere Formen der Verhaltensbeeinflussung scheitern. Denn während die Fahrer:innen in diesem Beispiel die Tempolimit-Schilder nicht wahrnehmen, weil sie von Chicagos Skyline abgelenkt sind, setzt Nudging auf vorbewusste Mechanismen und beeinflusst mithilfe einer optischen Täuschung das Fahrverhalten.

Das hier unterstellte Zusammenspiel der Wahrnehmungsapparate der Fahrer:innen und der weißen Streifen auf dem Boden ist zunächst nicht viel mehr als eine schöne Geschichte, deren Wahrheitsgehalt schwer überprüft werden kann. Ich habe im vorigen Kapitel deutlich gemacht, dass Umwelten nicht direkt beobachtbar sind, sondern durch aufwändige experimentelle Verfahren rekonstruiert werden müssen. Deshalb sind wir als Publikum dieser Geschichte darauf angewiesen, dass uns irgendwie gezeigt wird, dass die weißen Streifen am Lake Shore Drive wie intendiert Teil der Umwelt der Fahrer:innen werden. Dass die veränderte Umgebung mit Blick auf das

Bezugsproblem der hohen Zahl von Unfällen etwas ausrichtet, geht erst aus dem direkt darauffolgenden Absatz hervor, den ich im vorigen Kapitel noch nicht zitiert hatte. Tatsächlich scheint es so zu sein, dass die weißen Streifen die Zahl der Unfälle an diesem Straßenabschnitt drastisch reduziert haben:

Until the recent release of data by the Chicago Department of Transportation, only anecdotal accounts provided any indication of how effective the lines had been in preventing accidents. According to an analysis conducted by city traffic engineers, there were 36% fewer crashes in the six months after the lines were painted compared to the same sixmonth period the year before. (Thaler/Sunstein/Balz 2013: 433f.)

Erst durch diesen Absatz wird die Passung von Problem und Lösung validiert, indem ›harte‹, also daten- und zahlenbasierte Evidenz darüber angeführt wird, dass die Implementierung der weißen Streifen tatsächlich einen Effekt hatte. Wenn wir zunächst einen Blick auf die Gliederung der beiden zuletzt zitierten Absätze werfen, fällt auf, dass im ersten Absatz die Beschreibung von Problem und Lösung gemeinsam erfolgt, und die Evidenz dann einen eigenen Absatz bekommt. Problem und Lösung treten gemeinsam auf, die Evidenz gesellt sich zu ihnen und untermauert, dass sie ›wirklich‹ zueinander passen.

Ich habe im vorigen Kapitel die Ko-Konstitution von Problem und Lösung beschrieben, und in der Darstellung des Lake-Shore-Drive-Nudges sehen wir ein eindrückliches Beispiel dafür, wie zwischen Problem und Lösung im fertigen Nudge kein Blatt Papier mehr passt. Das Problem ist so beschrieben, dass es förmlich nach seiner Lösung in Form weißer Streifen auf dem Boden ruft, die ihrerseits als genau passende Antwort erscheinen. Und dennoch: Der unterstellte Wirkzusammenhang zwischen Problem und Lösung bliebe ohne die Evidenz reine Spekulation. Deshalb ist der zweite Absatz der Beschreibung so wichtig, in dem Daten beweisen, was anekdotische Evidenz zuvor nur vermuten konnte: Die weißen Streifen wirken. Nachdem sie implementiert wurden, gab es tatsächlich 36 Prozent weniger Unfälle. Erst als Trias, unter Hinzunahme der Evidenz, funktioniert die Geschichte und der Nudge als formatives Objekt. Erst dann überzeugt die Idee des sanften, des unbemerkten, des minimal-invasiven Anstupsens, mit dem große Wirkungen erzielt und Menschenleben gerettet werden können.

Die Sanftheit der Verhaltensbeeinflussung wird während der Herstellung von Nudges für die Praktiker:innen zum epistemischen Problem. Denn in obigem Beispiel wird auch deutlich, dass der intendierte Effekt nur zeitverzö-

gert und mit einem erheblichen Aufwand eines ganzen Teams von »city traffic engineers« bewiesen werden konnte. Wer sonst könnte Auskunft darüber geben, ob die weißen Streifen tatsächlich den erhofften Effekt erzeugen? Die Fahrer:innen selbst wohl kaum. Schließlich operieren Nudging- Interventionen so niedrigschwellig, dass sie Verhalten unbemerkt beeinflussen. Auch eine Beobachterin in der Situation, die sich beispielsweise am Straßenrand positioniert, könnte nur schwer erkennen, ob die Autos nun aufgrund der weißen Streifen abbremsen oder weil sie ohnehin auf eine Kurve zusteuern. Die kausale Rückführung der Geschwindigkeitsreduktion auf die weißen Streifen setzt aufwändigere Formen der Sichtbarmachung voraus, die in obigem Beispiel erst möglich werden, wenn Monate später Daten aus dem Chicago Department of Transportation veröffentlicht wurden. Erst mit Blick auf diese Daten wird ein Vorher-Nachher-Vergleich möglich.<sup>1</sup> Der Steuerungseffekt der weißen Streifen zeigt sich also nicht direkt in der Situation. Er wird nur erkennbar durch die Analyse von Daten und die »Visibilisierung von Mustern, die ›dahinter‹ liegen, also mit bloßem Auge nicht sichtbar sind« (Nassehi 2019: 79).

In einer soziologischen Sprache formuliert, stehen die Praktiker:innen hier vor der Herausforderung, ›den Akteursstatus ihrer Lösungen zu beweisen‹. Anders als in handlungstheoretischen Ansätzen, ist in praxeologischen Theorieströmungen der Status des sozialen Akteurs nicht auf intentional oder subjektiv sinnhaft handelnde Menschen beschränkt (vgl. Gugutzer 2017: 154). Auch Dinge können handeln, oder besser gesagt Teil von Handlungen sein. Statt theoretisch festzulegen, wer Handlungsfähigkeit besitzt und wer nur benutzt wird, arbeiten praxeologische Strömungen mit einem konsequentialistischen Handlungsbegriff (vgl. Gertenbach/Laux 2019: 121), der »Akteure bzw. Aktanten nicht über konkrete, vorab festgelegte Eigenschaften (wie etwa

---

1 In den Selbstbeschreibungen des Nudging wird meist die Rolle sogenannter randomisierter Kontrollstudien (RCT) zur Generierung von Evidenz hervorgehoben. Bei diesen epistemischen Verfahren werden zufällig gebildete Kontroll- und Behandlungsgruppen miteinander verglichen. So sollen externe Wirkfaktoren ausgeschlossen werden, weil sich beide Gruppen nur hinsichtlich der Frage unterscheiden, ob sie der Intervention ausgesetzt waren oder nicht. RCTs werden in den Selbstbeschreibungen häufig als ›gold standard‹ der Evidenz bezeichnet und anderen epistemischen Praktiken als überlegen dargestellt (vgl. Feitsma 2019: 35ff.). Praktisch betrachtet sind jedoch häufig die Bedingungen zur Randomisierung nicht gegeben, weshalb auch andere Verfahren der Evidenzproduktion Anwendung finden. Beispielsweise der Vorher-Nachher-Vergleich am Lake Shore Drive.

Interessen, Bewusstsein, Reflexivität u. a.) [definiert], sondern (konsequentialistisch) von ihren Wirkungen her« Gertenbach/Laux 2019: 32f.). Bruno Latour (2007) hat dies auf die berühmte Formel gebracht, nach der »jedes Ding, das eine gegebene Situation verändert, indem es *einen Unterschied macht*, ein Akteur [ist]« (123, Hervorhebung: T.S.).

Im Beispiel des Lake Shore Drives stehen die Praktiker:innen genau vor dem Problem, nachweisen zu müssen, dass ihre weißen Streifen einen Unterschied machen. Schließlich sind sie nicht zu Dekorationszwecken angebracht, sondern weil die Fahrer:innen zu schnell in die Kurven zu rasen drohen. Doch scheint es alles andere als trivial zu sein, den Effekt der weißen Streifen auch nachzuweisen. Er ist nicht in der Situation beobachtbar, sondern nur in Daten zu finden. Um diesen Sachverhalt noch deutlicher herauszustellen, möchte ich das Beispiel des Lake Shore Drives mit einem anderen Beispiel kontrastieren, in dem es ebenfalls darum geht, über veränderte Umgebungen die Geschwindigkeit von Autos zu reduzieren. Dieses Beispiel entspringt nicht den Nudging-Diskursen sondern soziologischen Debatten zu verteilter Handlungsmacht. Es ist Bruno Latours Beispiel des ›schlafenden Polizisten‹, der Bodenschwelle auf Straßen, welche Fahrer:innen dazu zwingt, abzubremsen:

[D]ie Bodenschwelle in einer verkehrsberuhigten Straße oder auf einem Campusgelände, um die Autofahrer zum Langsamfahren zu zwingen, die in Frankreich ›schlafender Gendarm‹ genannt wird. Durch die Schwelle wird das Ziel des Fahrers einer Übersetzung unterzogen. Sein ursprüngliches Ziel lautete: ›Fahre langsam, damit du keine Studenten gefährdest‹, die Übersetzung dagegen: ›Fahre langsam, damit deine Stoßdämpfer geschont werden.‹ [...] Durch die Vermittlung der Bodenschwelle verändert der Fahrer sein Verhalten: Aus der Moral fällt er ins Reich des Zwangs zurück. Doch von einem Beobachterstandpunkt aus betrachtet ist es gleichgültig, auf welchem Wege eine bestimmte Verhaltensweise erreicht wird. Von ihrem Fenster aus sieht die Rektorin der Universität, daß ihre Anordnung befolgt wird und die Autos abbremsen, und das reicht ihr. (Latour 2002a: 226f.)

Ich führe dieses Beispiel vor allem an, weil hier die Beobachtbarkeit des Steuerungseffektes völlig unproblematisch zu sein scheint. Der unterstellte Wirkzusammenhang, demzufolge die Fahrer:innen abbremsen müssen, um ihr Auto nicht zu beschädigen, ist so grob und so plausibel, dass keine Zweifel bezüglich seiner Wirkung aufkommen. Im Vergleich zur sanften Beeinflussung am Lake Shore Drive, wo erst Monate später ein Effekt nachweisbar ist, übt der

›schlafende Polizist‹ puren Zwang aus, an dem nichts Subtiles ist. Würden die Fahrer:innen nicht abbremsen, beschädigen sie sich und ihr Auto. Und würden sie nach dem Grund ihres Abbremsens befragt, könnten sie gewiss auf die Bodenschwelle verweisen.<sup>2</sup> Und auch die Rektorin der Universität, die Latour in seiner Beschreibung als Beobachterfigur erscheinen lässt, kann nach einem kurzen Blick aus dem Fenster zufrieden feststellen, dass die Autos ihre Geschwindigkeit reduzieren und der ›schlafende Polizist‹ seine Arbeit zuverlässig erledigt.

Im Vergleich zum Lake Shore Drive wirkt Latours Beispiel wie ein Ausflug in eine frühere Zeit. Schematisch lässt sich der ›schlafende Polizist‹ als eine Analogtechnik der Verhaltensbeeinflussung verstehen, die so etwas wie eine »Eins-zu-eins-Übertragung von Ursache und Wirkung, Signal und Reaktion, Steuerung und Umsetzung vorsieht« (Nassehi 2019: 20). Wird die Bodenschwelle errichtet, bremsen die Fahrer:innen ab, um Schaden abzuwehren. Täten sie es nicht, würde ihr Auto beschädigt. Dem steht die ›versiertere‹ Wirkungsweise der weißen Streifen gegenüber, bei denen Ursache und Wirkung, Signal und Reaktion, Steuerungsimpuls und Steuerungseffekt nur über Umwelten vermittelt auftreten und auf kompliziertere Weise miteinander in Beziehung stehen.

## Evidenz als Attraktor der Arbeit an Problem und Lösung

Für die Praktiker:innen liegt in der Sanftheit ihrer Form der Verhaltensbeeinflussung Fluch und Segen zugleich. Weniger Zwang auszuüben, mag wirksam und vielversprechend klingen. Doch weil die Effekte veränderter Umgebungen so flüchtig sind, ist eine Menge Arbeit nötig, um sie überhaupt sehen zu können. Deshalb spielen Fragen der Mess- und Beobachtbarkeit während der gesamten Herstellungsdauer von Nudges eine wichtige Rolle. In den Selbstbeschreibungen und in Phasenmodellen erscheint die Arbeit an der Evidenz als eine Phase, welche sich erst anschließt, nachdem Probleme analysiert, und Lösungen entwickelt und implementiert wurden. Doch mit Blick auf die Pra-

---

2 Während der Arbeit an dieser Studie hatte ich das Vergnügen, als Mitfahrer in einem Taxi durch einen Straßenabschnitt zu fahren, in dem eine Reihe schlafender Polizisten angebracht war. Die Fahrt war bis dahin eher schweigsam verlaufen. Doch die Bodenschwellen veranlassten den Fahrer dazu, mir seine Interpretation der Dinge mitzuteilen: »Schlecht für das Auto, aber gut für die Kinder.«

xis des Nudging wird deutlich, dass sich die Arbeit an der Evidenz über den gesamten Herstellungsprozess von Nudges erstreckt.

Was immer die Praktiker:innen tun, sie tun es (auch) mit Blick auf die Evidenz, die während der gesamten Arbeit am Nudge ihren Schatten vorauswirft und die Gestaltung von Problem und Lösung maßgeblich beeinflusst. Dementsprechend verstehe ich »Evidenz nicht als Repräsentation« (Ehlers/Zachmann 2019: 19) einer Wirkung, die durch veränderte Umgebungen erzeugt wurde und lediglich sichtbar werden muss. Evidenz erscheint vielmehr als ein spezifisches »Arrangement von Praktiken« (ebd.), das den gesamten Herstellungsprozess prägt. Was als Problem bearbeitet wird, wie die Lösung gestaltet wird, ist in entscheidendem Maße auch von der Anforderung bestimmt, dass Evidenz generiert werden muss.

Wie kann ich nun das Verhältnis der Evidenz zu den beiden anderen Teilobjekten des Nudges als formativem Objekt beschreiben? Ich habe im vorigen Kapitel herausgearbeitet, dass sich Problem und Lösung in einem ko-konstitutiven Verhältnis gemeinsam herausbilden. Die Evidenz erschien dabei als zusätzliche Anforderung, welche die miteinander verwobenen Problem- und Lösungskarrieren in eine bestimmte Richtung lenkt. Um ihre Wirkung auf einen Begriff zu bringen, bezeichne ich die Evidenz als Attraktor der Trajektorien von Problem und Lösung. Der Begriff ›Attraktor‹ hat seinen Ursprung in der Theorie komplexer Systeme und wurde beispielsweise von Manuel de Landa (2013) in seinem Buch »Intensive Science and Virtual Philosophy« in den geistes- und sozialwissenschaftlichen Diskurs überführt. De Landa beschreibt die Wirkung eines Attraktors wie folgt:

[A] large number of different trajectories, starting their evolution at very different places in the manifold, may end up in exactly the same final state (the attractor), as long as all of them begin somewhere within the ›sphere of influence‹ of the attractor (the *basin of attraction*). Given that, in this sense, different trajectories may be attracted to the same final state, singularities are said to represent the inherent or intrinsic *long-term tendencies* of a system, the states which the system will spontaneously tend to adopt in the long run as long as it is not constrained by other forces. (Ebd.: 7)

Wenn nun die gesamte Herstellung von Nudges auf die Produktion von Evidenz hinausläuft, die gesamte Arbeit am Nudge in der ›Einflussosphäre‹ der Evidenz liegt und sich die Trajektorien von Problem und Lösung mit der Evidenz auf einen gemeinsamen Zustand hinbewegen, dann müsste es möglich

sein, zu jedem Zeitpunkt der Herstellung von Nudges die Wirkung der Evidenz auch zeigen zu können. Ich habe sie bisher immer wieder als anwesende Abwesende beschrieben und möchte nun an einigen Beispielen veranschaulichen, wie sie während der Herstellung sichtbar wird.

Schon zu Beginn der Herstellung von Nudges finden Einschätzungen statt, ob sich die angestrebten Steuerungseffekte auch messen lassen werden und ob Evidenz hergestellt werden kann. Weil die Praktiker:innen wissen, dass ein Nudge ohne Aussagen über seine Wirksamkeit nicht als solcher verkauft werden könnte, spielen Fragen der Mess- und Beobachtbarkeit schon zu Beginn eine Rolle. Dies lässt sich an folgendem Beispiel verdeutlichen, in dem ein Nudge bereits in einem sehr frühen Stadium als potenzielle Projektidee daraufhin geprüft wird, ob sich für ihn auch Evidenz beschaffen lässt:

Beim Wochenabschlussmeeting in entspannter Atmosphäre sitze ich mit den Mitarbeiter:innen beisammen und wir trinken gemütlich Bier. Peter erzählt von einer Projektidee: Wie könnte man Kinder dazu bringen, sich gesünder zu ernähren? Oft würden Familien mit Bildungsangeboten adressiert, in der Hoffnung auf diese Weise das Ernährungsverhalten der Kinder zu ändern. Aber ihm schwebt etwas anderes vor: Was wäre, wenn man Kindern einfach im Klassenzimmer geschnittenes Obst und Gemüse vorsetzen würde und dann zeigen könnte, dass sie tatsächlich mehr davon essen. Mit Blick auf das Ernährungsproblem, bei dem sich Bildungsangebote als unwirksam erweisen, wären das doch überzeugende Ergebnisse! Auf eine unkonventionelle aber wirksame Weise wäre das Problem adressiert. Laura schaltet sich ein. Das klinge zwar gut, meint sie, aber wie wolle man hier denn eine Baselinemessung machen? Aktuell gebe es ja weder Obst noch Gemüse im Klassenzimmer, so dass es schwer werde, eine Veränderung des Konsums zu zeigen. Ihr Kollege stimmt ihr zu. Er müsse wohl noch weiter darüber nachdenken. (18.10.19)

In dieser frühen Phase, in welcher der Nudge nur als Idee existiert, wird er schon darauf geprüft, ob hier Evidenz hergestellt werden kann. In der Beschreibung seiner Projektidee fokussiert sich der Mitarbeiter auf den Problem- und Lösungsaspekt des Nudges. Doch sofort springt seine Kollegin ein und moniert, dass das Projekt unter Messgesichtspunkten so nicht durchführbar wäre. Um eine Verhaltensänderung zeigen zu können, braucht es eine Messung vor der Implementierung der Lösung, von der aus dann per Vergleich eine Erhöhung des Obst- und Gemüsekonsums festgestellt werden könnte. Das Problem, an dem ihr Kollege arbeiten möchte, scheint sich also

vor der Transformation in eine Variable zu sträuben, von der ich im Problemkapitel berichtet habe. Es hapert an der Evidenz. Wenn der Mitarbeiter seine Projektidee nun trotzdem verfolgen würde, verliese er die Einflussosphäre des Evidenz-Attraktors. Natürlich könnte er den Schüler:innen trotzdem Obst und Gemüse vorsetzen. Nur wäre es ihm nicht möglich, prozentuale Angaben über den Zuwachs an Obst- und Gemüsekonsum zu generieren.

Im weiteren Verlauf der Herstellung, solange noch keine Gewissheit darüber besteht, ob der gewünschte Effekt auch tatsächlich eintritt, herrscht oft große Anspannung unter den Praktiker:innen. Oft konnte ich Momente des Bangens und Hoffens miterleben, in denen noch nicht klar war, ob Evidenz entsteht und ob es die ›richtige‹ Evidenz sein wird. Denn zunächst gilt es eine Reihe von Bedingungen sicherzustellen, um überhaupt valide und statistisch signifikante Aussagen treffen zu können. Das sind die basalen methodischen Anforderungen an die Evidenz: Daten müssen im richtigen Format, in ausreichender Fallzahl und passend zum Erhebungszeitraum vorliegen, so wie es im Experiment vorgesehen ist. Darüber hinaus muss es dann aber auch die ›richtige‹ Evidenz sein, die eine Veränderung der gemessenen Variablen in die gewünschte Richtung nachweist. Nur dann können die Messungen in veränderten Umgebungen als Verbesserung interpretiert werden. Es gehört zwar zum Selbstverständnis des Nudging, auch Ergebnisse zu publizieren, wenn kein Effekt nachgewiesen werden konnte. Doch auf Dauer könnte Nudging nicht bestehen, wenn nicht Verbesserungen bewiesen würden.

Um der Ungewissheit in Bezug auf die Evidenz zu begegnen, versuchen die Praktiker:innen während der Herstellung, einen Blick auf vorläufige Ergebnisse zu erhaschen:

In der Statusbesprechung fragt Joseph, ob Judith schon die Möglichkeit hatte, einen Blick auf die Daten des laufenden Experiments zu werfen: »It would be such a shame to do this in a month and then see nothing!« Für Judith bedeutet das zusätzliche Arbeit. Denn sie muss dazu ihre Kontaktperson beim Projektpartner kontaktieren, die dort für die Daten zuständig ist, und um einen vorzeitigen Datenausgang bitten. Deshalb zögert sie. Maria pflichtet Joseph bei und sagt, dass der Aufwand gerechtfertigt ist: »Just so that we can sleep at night. Because like if we know that: Ok, it works, and everything is fine. Then we can relax a little bit«. (Notiz 11.11.19)

Entsprechend groß ist dann oft die Erleichterung, wenn sich die gewünschten Effekte abzuzeichnen beginnen und die Evidenz zu Problem und Lösung zu

passen scheint. Zwei Wochen später kommt die Mitarbeiterin mit guten Nachrichten in die Statusbesprechung:

I think the data says pretty much what we want it to say. I think we pretty much have the answers that we need. I haven't really made like a deep analysis. I just looked at it quickly. (Notiz 25.11.19)

Und wieder etwas später, stellt sich dann die Entspannung ein, wenn die Ergebnisse endlich da sind, und aus der anwesenden Abwesenden eine Anwesende wird. Jetzt ist bewiesen, dass die entwickelte Lösung wie erhofft wirkt:

Während ich ins Büro komme, schließt Joseph die Datenanalyse ab und präsentiert mir stolz seine Ergebnisse: »Fresh out of the oven«, sagt er, »like we thought it would be.« Er zeigt mir drei Balken, von denen der dritte viel höher ist als die anderen beiden. (Notiz 13.12.19)

Wenn am Ende eines Projekts Evidenz in Form eines Balkendiagramms vorliegt, mit dem die Wirksamkeit des Nudges untermauert werden kann, kommt die Herstellung des Nudges zu ihrem Ende. Entlang des Antizipierens, des Hoffens auf und des Bangens um die Evidenz, können die Trajektorien von Problem und Lösung nachverfolgt werden, bis sich unter den Praktiker:innen Erleichterung einstellt, weil Evidenz entstanden ist und der Nudge als formatives Objekt seine dreigliedrige Gestalt gewinnt.

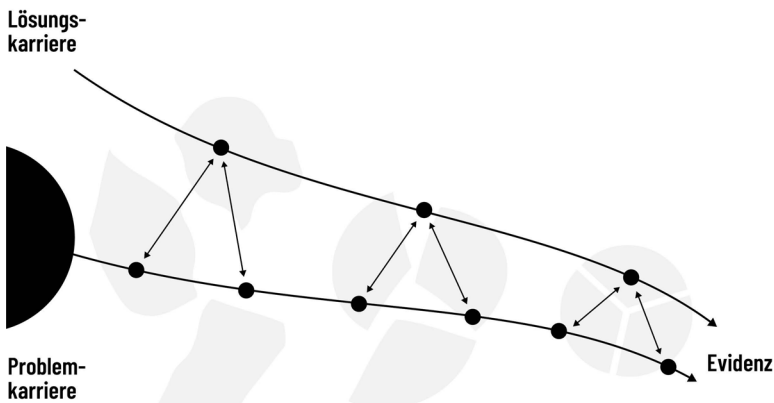
Ich werde mich im weiteren Verlauf dieses Kapitels nun auf drei unterschiedlichen Ebenen mit der Arbeit an der Evidenz beschäftigen. Während ich in den vorigen Kapiteln die Karrieren der jeweiligen Teilobjekte rekonstruiert habe, und mich je einem Abschnitt der Problem- oder Lösungskarriere widmete, muss ich hier im Evidenzkapitel anders verfahren. Denn Problem- und Lösungskarriere ergeben in ihrem Zusammenspiel die Evidenz, so dass ich nicht gleichartig von *einer* Evidenzkarriere ausgehen kann. Die Gliederung des restlichen Evidenzkapitels ist also weniger von der Empirie als von drei konzeptuellen Überlegungen bestimmt: Als erstes werde ich unter dem Stichwort *Laboratisieren* die Inhalte der beiden vorhergehenden Kapitel als Laboratisierungsprozess und als Errichtung eines Experimentalsystems interpretieren. Ich blicke also in diesem Abschnitt auf die gesamten Verläufe der Problem- und Lösungskarrieren zurück. Im Anschluss werde ich unter dem Begriff *Verdoppeln* die Eigenschaften des sanften Steuerungseffekt von Nudges weiter herausarbeiten, der aufwändig sichtbar gemacht und in Daten gesucht werden muss.

Unter dem Begriff *Wiederholen* stelle ich dann einige Überlegungen und Prognosen zur weiteren Durchsetzung des Nudging an.

## V.1 Laboratisieren

Beim Blick zurück auf die vorigen beiden Kapitel lässt sich nun erkennen, dass die gesamte Arbeit an Problem und Lösung auf die Produktion von Evidenz hinausläuft. Während Evidenz zu Beginn der Herstellung von Nudges qua Negation wirksam ist, indem sie vor allem dafür sorgt, dass manche möglichen Wege nicht eingeschlagen werden, manifestiert sie sich im Verlauf der Herstellung von Nudges immer weiter, bis sie am Ende als Balkendiagramm greifbar und vorzeigbar ist. Um das Werden der Evidenz zu verdeutlichen, komme ich nun erneut auf die Abbildung zurück, die ich im bisherigen Verlauf dieser Studie entwickelt habe. Darin erkennen wir nun, dass sich Problem und Lösung nicht nur ko-konstituieren, sondern dass sich ihre verwobenen Karrieren auch gemeinsam auf die Evidenz hin entwickeln, die wiederum nur im Zusammenspiel von Problem und Lösung entstehen kann. Die Evidenz wirkt als Attraktor der miteinander verwobenen Problem- und Lösungskarrieren.

Abbildung V.1 Die Evidenz als Attraktor von Problem- und Lösungskarriere



Das Problem durchläuft im unteren Pfeil der Abbildung die Transformationen zum *Fluss*, zur *Engstelle*, zum *Verhalten*, zur *Entscheidung*, zur *Variablen* und zur *Differenz*. Dabei verengt sich die Problemkarriere trichterförmig auf immer kleinere und immer stärker lokalisierte Einheiten, bis das Problem am Ende als messbare und gemessene Variable fixiert ist. Zeitgleich kristallisiert sich im oberen Pfeil entlang von *Umgebung*, *Umwelt* und *Experimentalaufbau* die Lösung heraus, die um das Problem herum existiert, und an die immer neue Anforderungen hinzukommen, bis sie am Ende in doppelter Hinsicht unter Kontrolle gebracht ist: Sie muss sowohl veränderbar sein als auch Beobachtungen und Messungen ermöglichen. Sind Problem und Lösung derart in ein Verhältnis gebracht, kann Evidenz entstehen.

### Der Nudge als Experimentalsystem

Das so beschriebene Zusammenwirken von Problem und Lösung möchte ich nun mithilfe von Begriffen des Wissenschaftshistorikers Hans-Jörg Rheinberger als Herausbildung eines Experimentalsystems beschreiben. Denn das Verhältnis von umgebenem Problem und umgebender Lösung erinnert in seiner Relationalität an das, was Rheinberger »Experimentalsysteme« (Rheinberger 2006) nennt, und in denen umgebene epistemische Dinge und umgebende technische Dinge miteinander in Beziehung stehen.<sup>3</sup> »Betrachtet man Experimentalsysteme etwas näher«, schreibt Rheinberger, »so wird man feststellen, dass in ihnen zwei verschiedene, jedoch voneinander nicht trennbare Strukturen ineinandergreifen« (ebd.: 27) – zwei Strukturen, deren Verwobenheit ich in den vorigen beiden Kapitel herausgearbeitet habe. Ersterer nennt Rheinberger die epistemischen Dinge, also jene »Dinge, denen die Anstrengung des Wissens gilt« (ebd.) und die in einer für sie »charakteristischen, irreduziblen Verschwommenheit und Vagheit« (ebd.) erscheinen. Sie sind eingefasst in weitere technische Dinge, also in »stabile Umgebungen, die man als Experimentalbedingungen [...] bezeichnen kann« (ebd.: 29). Dazu zählt Rheinberger vor allem »Instrumente [und] Aufzeichnungsapparaturen« (ebd.), mit denen auf epistemische Dinge eingewirkt werden kann und mit deren Hilfe sich messen lässt. Die technischen Dinge determinieren die epistemischen Dinge »in doppelter Hinsicht: Sie bilden ihre Umgebung und lassen sie so erst als solche hervortreten, sie begrenzen sie aber auch

3 Die Formulierung von Umgebenem und Umgebendem geht auf die Studie »Epistemologien des Umgebens« von Florian Sprenger (2019) zurück.

und schränken sie ein.« (Ebd.) Diesen Aspekt habe ich bisher als Spannung zwischen Steuerungsambition und Messanspruch thematisiert.

Mit diesen Begriffen kann ich nun den Inhalt der ersten beiden Kapitel in einem neuen Licht betrachten. Das Problem, welches ich entlang seiner Transformationen beschrieb, erscheint nun als im Werden begriffenes epistemisches Ding. Als *Fluss*, als *Engstelle*, als *Verhalten*, als *Entscheidung*, als *Variable* und als *Differenz* ist es genau das, worauf die Forschungsbemühungen der Praktiker:innen gerichtet sind, das sie zu fassen versuchen, und das, »paradox gesagt, das [verkörpert], was man noch nicht weiß« (Rheinberger 2006: 28). Egal wie deutlich das Bezugsproblem am Anfang erscheint, jede seiner Transformation in der Problemkarriere ist kontingent und hängt davon ab, was vorher war und was die Umgebung jeweils bereithält. Das zuvor beschriebene Bangen, ob sich tatsächlich die erhofften Verhaltensänderungen einstellen, entspricht genau diesem ungewissen Bemühen, epistemische Dinge zu fassen zu bekommen.

Die Lösung wiederum, welche ich als *Umgebung*, als *Umwelt* und als *Experimentalaufbau* beschrieb, kann nun als technisches Ding verstanden werden. Es fasst Probleme ein, umgibt sie und muss – wie ich im vorigen Kapitel argumentiert habe – in doppelter Weise unter Kontrolle gebracht werden, genau wie es Rheinberger mit der Nennung von Instrumenten und Aufzeichnungsapparaturen impliziert. Denken wir nur zurück an die Schilderungen eines Informanten im ersten Kapitel, der von seiner »fake canteen« berichtete: Ein Ort, wo er verändernd in das »environment« eingreifen konnte, um auf Ernährungsverhalten einzuwirken, und dabei »all these different measuring tools« zur Verfügung hatte.

Der Nudge als Experimentalsystem und die Arbeit am Nudge als dessen Errichtung – so verstanden ist Nudging ein Regieren durch Verhaltensexperimente. Damit kann ich nun auch die Verbindung des Nudging zur Verhaltensökonomik anders beschreiben. Denn in der Einleitung dieser Studie ging es darum, dass Nudging aus einer verhaltensökonomischen Kritik am »Homo Oeconomicus« entstanden ist. Damit machte ich die Verbindung des Nudging zu diesem Forschungsgebiet entlang des spezifischen Menschenbildes deutlich, das in der Verhaltensökonomik erarbeitet wurde und den Nudging-Interventionen zugrunde liegt. Diese Interpretationsweise einer politischen Anthropologie kann ich nun durch eine Perspektive ergänzen, welche nicht auf Menschenbilder, sondern auf epistemische Praktiken ihren Fokus legt. Nudging ist demnach eine Regierungstechnologie, die der experimentellen Forschung im Labor entstammt. Deshalb ist sie dazu verdammt, die Welt durch

die Brille des Experimentators zu betrachten und jedes Problem, das sich ihr stellt, zunächst in jene Entitäten zu transformieren, welche sie als epistemische Objekte aus Laborexperimenten kennt – so wie »im Prozess des Experiments [...] eine spezifische wissenschaftlich und technisch durchstrukturierte Umwelt für den wissenschaftlichen Gegenstand geschaffen [wird], welchen die Forschenden untersuchen wollen« (Haus 2021: 2).

Rheinberger weist darauf hin, dass sich Experimente immer nur als Teile von »Reihen von Experimenten« (Rheinberger 2006: 30) verstehen lassen, die als solche zwar »in einigen Richtungen klarer, aber zwangsläufig auch weniger unabhängig [werden], weil sie sich zunehmend auf eine Hierarchie bereits etablierter, aufeinander bezogener und abgestimmter Prozeduren stützen« (ebd.). Diesen Gedanken führt er dann, Ludwik Fleck zitierend, weiter aus und schreibt:

Ist ein Gebiet bereits so ausgearbeitet, daß die Schlußmöglichkeiten auf Existenz oder Nichtexistenz, eventuell auf quantitatives Feststellen begrenzt sind, so werden die Experimente immer klarer, sie sind aber nicht mehr selbstständig, da sie vom System früherer Experimente und Entscheidungen geschleppt werden. (Fleck 1980: 114, zit. nach Rheinberger 2006, 30)

Nudging als Regieren durch Verhaltensexperimente wird somit vom System der früheren Experimente und Entscheidungen verhaltensökonomischer Forschung geschleppt. Es wurde im ökonomischen Verhaltenslabor geboren, das »die Komplexität und Verwobenheit der alltäglichen Wirklichkeit [...] auf ein analytisch fruchtbares Maß reduzier[t]« (Haus 2021: 2). Weil es sich aber nun Problemen außerhalb des Labors zu stellen versucht, kann es dorthin nicht mehr zurück. Stattdessen versucht es, die Welt zum Labor zu machen und reinszeniert fortwährend einen Zustand, in dem die Dinge kontrollierbar sind.

## Die Herstellung des Nudges als Laborisierungsprozess

Wie gelingt das? Wie wird die Welt zum Labor? Wie kann die Komplexität und Verwobenheit der Welt so reduziert werden, dass kausale Eingriffe möglich werden? Rheinbergers Begriffe des Experimentalsystems, der epistemischen und der technischen Dinge entstammen der Untersuchung biologischer Grundlagenforschung. Dort gibt es immer schon Labore, deren Mauern nach außen hin abschotten. Nudging hingegen findet nicht in Laboren statt, und

beschränkt sich auch nicht auf die Schaffung neuen Wissens. Der Anspruch lautet vielmehr, direkt in die Welt zu wirken, sie zu verändern und Probleme zu lösen. Die Weltausschnitte, in die eingegriffen wird, müssen deshalb – zumindest vorübergehend – zu Laboren gemacht, also *laboratisiert* werden (vgl. Knorr-Cetina 1988: 87f; Callon/Lascoumes/Barthe 2009: 65ff; Guggenheim 2012).

Michael Guggenheim denkt Laborisierung als einen Grenzziehungsprozess, bei dem das vernachlässigbare Außen vom relevanten Innen unterschieden wird, und dessen Resultat ein Labor ist, in dem Experimente stattfinden können:

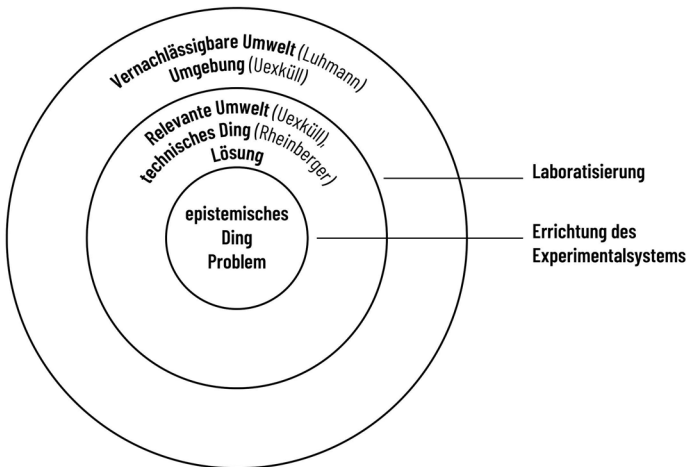
The laboratory is the result of a procedure that separates between an outside, an environment that is considered negligible for some epistemic claim or technological invention, and an inside, a (partly) controlled environment that is considered relevant for this claim or invention. The lab is not so much a closed space, but a procedure that often results in a space with the properties to separate controlled inside from uncontrolled outside. (Guggenheim 2012: 101)

Laborisierungsprozesse erfolgen Guggenheim zufolge also nicht nur, indem Mauern errichtet werden, die gegen das vernachlässigbare Außen abschotten. Auf einer abstrakteren Ebene können Labore auch durch Unterscheidungen erfolgen, die festlegen, was beobachtet und was als irrelevant erachtet wird. Interessant ist nun, dass Guggenheim eine Unterscheidung thematisiert, die im Vergleich zu Rheinbergers Unterscheidung von epistemischen und technischen Dingen ›weiter außen‹ liegt. Guggenheim beginnt mit der Unterscheidung zwischen einer *äußeren*, vernachlässigbaren Umwelt (»an environment that is considered negligible«) und einer *inneren*, (teilweise) kontrollierten Umwelt, die ihren Status je mit Blick auf den »epistemic claim or technological invention« gewinnt. Diese *innere* Umwelt kann mit Rheinberger, der die Abschottung nach außen immer schon voraussetzen kann, als technisches Ding verstanden werden, die ihren Status durch ihre Beziehung zum epistemischen Ding erhält.

Damit können zwei unterschiedliche Umwelten unterschieden werden: *Erstens* die ›äußere‹ Umwelt, die als vernachlässigbar gilt, weil ihr mit Blick auf das epistemische Ding keine Bedeutung zugesprochen wird. Bei der Arbeit am Nudge sind dies all die Dinge jenseits der Umgebung, die womöglich von einem Beobachter für das Problem als relevant erachtet werden könnten, die

aber aus Nudging-Perspektive keine Rolle spielen. Denken wir nur zurück an meinen Versuch im Lösungskapitel, das Problem der zu vollen Kantine durch Eingriffe in den Tagesablauf der Behörde zu beheben. Diese äußere Umwelt ist eine Umwelt im Verständnis der Systemtheorie Niklas Luhmanns (vgl. 1987: 242ff.), die als nicht relevant erachtet und damit vom relevanten Innen unterschieden wird, hierbei aber konstitutive Voraussetzung für jenes Innen ist, das nur in Abgrenzung zum Außen seine Identität gewinnt. *Zweitens* gibt es die ›innere‹ Umwelt, die als technisches Ding in Relation mit dem epistemischen Ding steht. Bei der Arbeit am Nudge sind dies all jene Aspekte der Umgebung, die am Problem eine Veränderung bewirken können. Dies ist Umwelt im Sinne Uexkülls, deren Besonderheiten ich im vorigen Kapitel herausgearbeitet habe.

Abbildung V.2 Laboratisierung und Errichtung des Experimentalsystems



Diese doppelte Grenzziehung habe ich in Anlehnung an (Pohl/Rupp 2021: 27) in Abbildung V.2. veranschaulicht. Bei der Herstellung von Nudges werden diese Grenzen gezogen durch eine spezifische Weise, auf die Welt zu blicken. Wenn die Welt durch die ›experimentelle Brille‹ der Praktiker:innen betrachtet wird, so gehen damit Festlegungen und Unterscheidungen einher. Jede Festlegung auf eine Entscheidungsarchitektur stellt auf bestimmte Entscheidungen und bestimmte relevante Umgebungen scharf, während alles andere aus dem

Blick gerät. Je nachdem, was von den Praktiker:innen gemessen wird, und zu welchen Aspekten der Umgebung sie die Messung in Beziehung setzen, ziehen sie die Grenzen zwischen vernachlässigbarer und relevanter Umwelt sowie zwischen relevanter Umwelt und Problem.

Es scheint also ein ›experimenteller Blick‹ zu sein, durch den diese doppelte Grenzziehung zustande kommt. Zugleich gibt es immer auch bestimmte Situationen, die sich besonders dazu eignen, laboratisiert zu werden. Das wurde mir in folgendem Beispiel deutlich. Bei unserer Exkursion zum Flughafen blieben wir für längere Zeit vor einem Gang stehen, den Passagiere entlanggingen, wenn sie nach der Landung ihr Gepäck abgeholt hatten und sich dann auf den Weg aus dem Landebereich machten. Am Ende dieses Ganges waren zwei automatische Türen, hinter denen typischerweise die Personen auf sie warteten, die sie abholten.

Wir stehen etwas abseits mit Blick in den Gang, an dessen Ende sich die zwei automatischen Türen befinden. Peter berichtet, dass hier das Problem aufgetreten sei, dass fast 90 Prozent der Passagiere die rechte Tür benutzt hätten. Auf der anderen Seite habe das immer wieder zu unnötigem Gedränge geführt. Während er ausführt, dass sie mit verschiedenen Markierungen auf dem Boden herumprobiert hätten, und dass schon die Farbwahl dieser Markierungen teilweise zu einem Unterschied von 8 Prozent geführt hätte, blicke ich in den Gang und warte immer gespannt darauf, welche Tür die entlangkommenden Passagiere durchschreiten. Plötzlich kann ich den Reiz des Nudging nachvollziehen und fühle mich zurückerinnert in den Physikunterricht, als wir das Doppelspaltexperiment kennenlernten, in dem nun jeder neue Passagier ein weiteres Teilchen wäre. Einem Kollegen scheint es ähnlich zu gehen. Er zählt mit und gibt irgendwann den Zwischenstand durch: Aktuell steht es 6:0 für die rechte Tür. (Notiz 12.11.19)

Diese Situation erscheint, »as if an experimenter has designed a laboratory experiment within the world« (Morgan 2013: 346). Zugleich gehört aber auch eine laborisierende Form der Beobachtung dazu, die wir als Gruppe hier praktizieren: Der gut einsehbare Gang, zwei Türen an seinem Ende, und nach und nach hindurchfließende Passagiere ermöglichten uns, hier ein Experimentalsystem zu erkennen, das dazu einlud, Interventionen vorzunehmen und ihre Effekte zu beobachten. Das Problem, um das es sich hier drehte, war bereits heruntergebrochen auf eine ungünstige Häufigkeitsverteilung der Nutzung von zwei Türen. Als Lösung waren verschiedene Markierungen auf dem Boden im Gespräch, die anscheinend je nach Farbe unterschiedlich wirksam

waren. Alles andere spielte in diesem Moment keine Rolle. Die Welt bestand nur noch aus den Markierungen, aus der rechten und der linken Tür. Dieser Bereich des Flughafens war für uns ein Labor mit Experimentalsystem. Was dabei an Komplexität verloren ging, gewannen wir an anderer Stelle zurück: Die Möglichkeit der kausalen Einwirkung.

Egal ob sie sich nun von selbst dazu anbietet oder vor allem durch die Art der Betrachtung zum Laborexperiment wird, Laboratisierungsprozesse sind etwas, das nicht nur *in der Welt*, sondern auch *an der Welt* geschieht. Nortje Marres und David Stark (2020) nehmen diesen Unterschied zum Anlass, das Forschungsprogramm einer »new sociology of testing« auszusrufen: »[T]esting and experimentalism«, so schreiben sie, »have been extended into distinctively social environments, so that it is now society and social life itself that is subject to the technological regime of testing.« (ebd.: 427) Diesen Übergang bringen sie auf die pointierte Formel eines Übergangs von »tests in settings« (ebd.: 437) zu »testing settings« (ebd.). Von Laboratisierung zu sprechen bedeutet in diesem Kontext nicht, dass die Gesellschaft insgesamt zu einem Labor wird, wie es beispielsweise in den 1990er Jahren im Kontext von Risikotechnologien und Ereignissen wie der Tschernobyl-Katastrophe diskutiert wurde (vgl. Krohn/Weyer 1990). Stattdessen ist damit zum Ausdruck gebracht, dass einzelne Weltausschnitte vorübergehend laboratisiert werden: »[A]t different spots laboratories are implanted that frame and preformat possible actions.« (Callon/Lascoumes/Barthe 2009: 67) Immer dann, wenn meine Informant:innen Situationen finden, die sie beobachten und überwachen können, und in deren Umgebungen sie verändernd eingreifen können, sind für sie die Bedingungen zur Laboratisierung gegeben.

Dieser neue Experimentalismus ist Marres und Stark zufolge dort am stärksten ausgeprägt, wo durch digitale Technologien der Vollzug des Sozialen und seine Messung ineinander fallen und somit potenziell jede Aktivität eine Spur in Form von Daten hinterlässt: »In the digital society today, the measurement of social life is performed as part of the conduct of social life, as every click and retweet generates *at the same time* a social action and a data point.« (Marres/Stark 2020: 436) Anders als im Flughafenbeispiel, wo ein Mitarbeiter zählt, und das Labor verschwinden würde, wenn er damit aufhört, ermöglichen digitale Umwelten als »total test environments« (ebd.) ein permanentes Experimentieren. So argumentiert auch Rainer Mühlhoff, wenn er davon spricht, dass »[d]urch den fast flächendeckenden Einsatz von Tracking-Techniken [...] das Internet [...] zu einem behaviorellen Echtzeitlabor« (Mühlhoff 2018: 562) werde.

In der Herstellungspraxis konnte ich die Tendenz beobachten, sich genau solche digitalen Umgebungen zu suchen, in denen Nudging leichter durchführbar ist. Das wurde im Lösungskapitel bereits deutlich, wo sich die Idee der Energiesparflyer, die bei Hausbesichtigungen überreicht werden sollten zur Verlinkung zweier Webseiten weiterentwickelte (siehe Kapitel IV.3). Dort argumentierte ich, dass die digitale Version leichter zu kontrollieren ist als die analoge. Dennoch lässt sich auch aus analogen Beispielen für uns als Beobachter:innen des Nudging sehr viel lernen. Denn alles was im Digitalen automatisiert stattfindet, muss außerhalb der total test environments mühsam erarbeitet werden. Deshalb widme ich mich im nächsten Abschnitt dem *Verdoppeln*, wo es darum geht, die Daten herzustellen, in denen sich die Steuerungseffekte von Nudges erst finden lassen.

## V.2 Verdoppeln

Am Beginn dieses Kapitels habe ich argumentiert, dass die Sanftheit des Steuerungseffektes von Nudges für die Praktiker:innen zum epistemischen Problem wird. Weil die Verhaltensbeeinflussung so subtil ist, dass sie unterhalb der Bewusstseinschwelle der Beeinflussten erfolgt, braucht es experimentelle Technologien der Sichtbarmachung, um den kleinen Stupser mit seiner großen Wirkung zu verbinden. Die Steuerungseffekte der Umgestaltung von Umgebungen sind mit dem bloßen Auge und mit Blick auf den Einzelfall nicht erkennbar. Denn statt Verhalten direkt zu beeinflussen, werden hier *Verhaltenswahrscheinlichkeiten* modifiziert. Die Veränderung von Umgebungen mag zwar einen Eingriff in die Welt darstellen, doch ihr Effekt zeigt sich nur im Bereich ihrer datenförmigen Verdoppelung (vgl. Luhmann 2002: 58; Nassehi 2019: 108ff.). Deshalb stehen die Praktiker:innen vor der Herausforderung, den Effekt »den Sinnen zugänglich, *sinnfällig* zu machen« (Rheinberger 2021: 17).

Luhmann (2002) bezeichnet »Realitätsverdoppelung« als eine Operation bei der »[i]rgendwelchen Dingen oder Ereignissen [...] eine besondere Bedeutung verliehen [wird], die sie aus der gewöhnlichen Welt (in der sie zugänglich bleiben) herausnimmt und mit einer besonderen »Aura«, mit besonderen Referenzkreisen ausstattet« (Luhmann 2002: 58). Damit ist Verdoppelung mehr als die bloße Abbildung von etwas, denn sie ermöglicht Anschlussoperationen, die ohne sie nicht denkbar wären. Mit dieser Beschreibung ist noch keine Festlegung auf die Art der Verdoppelung vorgenommen. Welche Ereignisse aus der

gewöhnlichen Welt herausgenommen werden und in welchem Medium die Verdoppelung erfolgt, kann ganz unterschiedlich sein. So sind das Spiel, die Kunst oder die Religion Luhmann zufolge als je spezifische Verdoppelungen Verwandte der statistischen Analyse (vgl. ebd.). Auch die Schrift ist eine Verdoppelung der Welt, welche bestimmte Anschlussoperationen ermöglicht, die sonst nicht möglich wären (vgl. Nassehi 2019: 113). Die datenförmige Verdoppelung der Welt, welche für das Verständnis von Nudging zentral ist, ist im Grunde nur ein anderer Begriff für das, was Armin Nassehi als Digitalisierung beschreibt: »[D]ie Abbildung der Welt in Form von zählbaren Einheiten und die Codierung von analogen Sachverhalten durch diskrete Formen.« (ebd.: 108)

Durch die datenförmige Verdoppelung der Welt wird der Vollzug des Sozialen in Datensätzen gespeichert, die das Geschehen hochselektiv repräsentieren und damit Analyseverfahren in bisher ungekanntem Ausmaß ermöglichen. Nassehi (2019) argumentiert in seiner »Theorie der digitalen Gesellschaft«, dass im Bereich der datenförmig verdoppelten Welt »nützliche und brauchbare Muster« (ebd.: 80) erkennbar werden »mit denen man digital sieht, was analog verborgen bleibt« (ebd.: 50). Darin liegt eine Asymmetrie begründet, zwischen denen, die die Daten haben, und jenen, deren Verhalten darin verdoppelt ist. Gesa Lindemann betont in diesem Zusammenhang: »Wenn das ›Leben‹ zu Daten verarbeitet worden ist, lassen sich in diesen Mustern erkennen, die diejenigen, die das ›Leben‹ leben, gar nicht sehen können.« (Lindemann 2020: 304) Als ich weiter oben im Anschluss an Goffman Nudging-Situationen als »environments of [unilateral, T.S.] monitoring possibilities« (Goffman 1964: 135) beschrieb, war damit auch dieser Aspekt gemeint.

Die datenförmige Verdoppelung der Welt ist in der Praxis mitunter ein sehr langsamer Prozess. Im Beispiel des Chicagoer Lake Shore Drives musste beispielsweise über sechs Monate auf die Veröffentlichung von Daten gewartet werden, um den Effekt der Implementierung von weißen Streifen ermitteln zu können. Weil er nur in den Daten zu finden ist, tritt der Effekt einer Nudging-Intervention also mit einer gewissen Latenz<sup>4</sup> auf, wobei Latenz verstanden wird als die »Zeit zwischen einem Ereignis und der darauf folgenden Reaktion« (Duden 2022b). In der Praxis der Herstellung von Nudges kann diese Verzögerungszeit sehr groß ausfallen. Zum Beispiel im Energieprojekt, das ich bereits in den vorigen Kapiteln thematisierte. Dort wurde neben der Implementierung von Energiesparflyern bei Hausbesichtigungen (siehe Kapitel III.2) auch die Idee verfolgt, solche Flyer in Beratungsgesprächen bei Banken

---

4 Ich danke Dennis Eckhardt für diesen Begriffsvorschlag.

einzusetzen und zu überprüfen, ob dies einen Effekt auf die Renovierungsbereitschaft oder die Zahl der tatsächlichen Renovierungen habe. Anders als bei den Hausbesichtigungen, die sich zu einem digitalen Experiment weiterentwickelten, wurde es bei den Banken tatsächlich analog praktiziert – mit entsprechenden Schwierigkeiten:

Ich habe mich in der Küche mit Judith über das Energieprojekt unterhalten. Das Experiment mit den Banken sei diese Woche losgegangen, doch bisher habe es nur zwei Gespräche mit Kund:innen gegeben. Beide Gespräche seien aber für das Experiment nicht relevant gewesen, weil ihre Häuser jeweils kein Energielabel gehabt hätten. Weil in den Flyern die Informationen aus dem Energielabel dargestellt werden, habe es also auch keine Flyer gegeben. Judith meint, das Experiment werde sich wohl noch eine Weile hinziehen und man müsse einfach abwarten. Sie brauche mindestens 50 Beratungsgespräche, in denen ein Flyer gezeigt wurde, um einen Effekt sehen zu können. (Notiz 06.12.19)

An dem Problem meiner Informantin wird besonders deutlich, wie weit die Intervention und ihr Effekt auseinanderliegen können. Zeitlich sind sie weit voneinander entfernt, weil es lange dauern wird, ehe genügend Fallzahlen vorliegen, mit denen sich ein Effekt nachweisen lässt. Gleichzeitig lässt sich aber auch eine räumliche und soziale Distanz erkennen, weil meine Informantin das Experiment zwar geplant hat, mit seiner Durchführung jedoch nichts zu tun hat, und mir nun kaffeekochend im Büro davon berichtet, wie es um die Dinge steht. Keine der Personen, auf deren Verhalten sie Einfluss nimmt, wird sie je zu Gesicht bekommen. Ihre Art, Verhalten zu beeinflussen, ist »governing at a distance« (Rose/Miller 1992) im besten Wortsinn.

## Die Herstellung des Steuerungseffektes

Dass der Steuerungseffekt mit einer Latenz, also außerhalb der beeinflussten Situation und mit einer Zeitverzögerung auftritt, lässt sich anhand eines anderen Beispiels weiter veranschaulichen. Auch dieses Beispiel kennen wir bereits aus den vorigen Kapiteln: Um die Zahl der Stürze auf Treppen zu reduzieren, wurden auf dem Boden im Treppenhaus eines Bürogebäudes Aufkleber in Form einer verunglückten Person angebracht. Im Problemkapitel habe ich dargelegt, wie es nur durch den Verweis auf eine Unfallstatistik gelang, das Geschehen im Treppenhaus zu einem »echten«, also für die Praktiker:innen re-

levanten Problem zu machen (siehe Kapitel III.3). Im Lösungskapitel habe ich die Frage verfolgt, warum ausgerechnet Aufkleber in Form einer verunglückten Person als Lösung implementiert wurden (siehe Kapitel IV.2). Nun werde ich mich der Arbeit widmen, durch die der Effekt dieser Intervention ermittelt wurde. Dieses Beispiel ist für mich deshalb interessant, weil an ihm gezeigt werden kann, vor welche Herausforderungen die Praktiker:innen gestellt sind, wenn sie die Effekte ihrer Interventionen sichtbar machen wollen. Abseits der »total-test-environments«, von denen Marres und Stark (2020: 436) sprechen, muss die datenförmige Verdoppelung der Welt in Handarbeit erfolgen. Aus analytischer Perspektive stellt sich das jedoch als Glücksfall dar, weil hier Prozesse beobachtbar werden, die sonst »in die Maschinerie selbst verlegt« (Rheinberger 2021: 33) und damit verborgen wären.

Wenn über Aufkleber auf dem Boden das Risikobewusstsein der Treppennutzer:innen beeinflusst werden soll, stehen die Praktiker:innen vor der Herausforderung, den unterstellten Wirkzusammenhang beweisen zu müssen. Als epistemisches Objekt kann das Problem des fehlenden Risikobewusstseins nicht direkt beobachtet werden. Stattdessen gehen die Praktiker:innen davon aus, dass seine Ausprägung in der Zahl der Hände am Geländer sichtbare Spuren hinterlässt. Wenn sich wenige Leute am Geländer festhalten – so ihre Überlegung – herrscht im Treppenhaus ein niedriges Risikobewusstsein. Eine Erhöhung von Händen am Geländer könnte dann auf eine Erhöhung des Risikobewusstseins hindeuten. Die Praktiker:innen müssen sich also entlang solcher Spuren auf die Suche nach dem Steuerungseffekt machen. Doch wie Rheinberger betont, sind »experimentelle Spuren [...] vergängliche Dinge« (Rheinberger 2021: 29). Um sie analysieren zu können, müssen sie zunächst haltbar gemacht werden. Deshalb hat sich ein Mitarbeiter mit seinem Laptop auf dem Schoß im Treppenhaus positioniert, wo er unter anderem die Häufigkeit der Hände am Geländer ermittelt. Er füllt dabei eine Exceltabelle aus.

In den ersten Spalten sind die Wörter »Kontrol«, ein Datum, der Ort der Beobachtung sowie ein Zeitraum zu lesen. Der Mitarbeiter hat diese Tabelle schon so vorbereitet, dass diese Spalten ausgefüllt sind, bevor er mit der Beobachtung beginnt. Er hat geplant, wann er beobachten wird, und er weiß, dass dies die Daten der Kontrollgruppe sein werden. »Before implementing any nudges, we needed to know the employees' current behaviour while descending the stairs. Therefore, an initial observation of 1,275 people was performed on two different locations.« (Projektbericht) Wenn er nun damit beginnt, die Tabelle auszufüllen, bringt er das Problem als Variable hervor und

übersetzt das Geschehen im Treppenhaus in eine digitale Repräsentation (vgl. Nassehi 2019: 108). Er arbeitet an der Verdoppelung der Welt und betreibt eine »Transformation von flüchtigen Spuren in haltbare Daten« (Rheinberger 2021: 29).

Wenn der Mitarbeiter auf dem Sofa sitzend in das Treppenhaus blickt, sieht er alles Mögliche. Seine Tabelle gibt aber vor, wonach er Ausschau halten soll; sie zwingt ihn zu Selektionen und Unterscheidungen. Mit jeder neu auf der Treppe erscheinenden Person füllt er eine neue Zeile aus. Er tippt Werte in die Felder, abhängig davon, ob die Person nach unten (0) oder nach oben (1) geht (Spalte 5), ob sie ihm weiblich (0) oder männlich (1) erscheint (Spalte 6), ob sie das Geländer fasst oder nicht (Spalte 7), ob sie etwas in den Händen trägt (Spalte 8), ob sie abgelenkt erscheint (Spalte 9), Schuhe mit hohem Absatz trägt (Spalte 10) und ob sie in Eile ist oder nicht (Spalte 11). Damit bringt er »die spurenhafte Ereignisse [...] in eine Form [...], in der sie gespeichert und auch wieder abgerufen und damit verarbeitet« (Rheinberger 2021: 32) werden können.

Nach dieser ersten Messung kann der Mitarbeiter später ins umgestaltete Treppenhaus zurückkehren und erneut beobachten. Wenn die Aufkleber tatsächlich einen Effekt auf das Risikobewusstsein haben, so wird dies Spuren hinterlassen, die erneut in Daten übersetzt und ausgewertet werden können. Weder die Treppengehenden noch der Mitarbeiter im Treppenhaus, könnten direkt einen Effekt erkennen. Wenn es ihn gibt, dann zeigt er sich nur in den Daten. Und tatsächlich: Ihre Analyse ergibt später, dass die Zahl der Hände am Geländer durch das Anbringen der Aufkleber um 9,2 Prozent gestiegen ist. Diese Zahl ist es nun, die den kleinen Stupser mit seiner Wirkung verbindet und die Wirksamkeit der Regierungsanstrengung untermauert. Ohne die epistemische Arbeit im Treppenhaus wäre der Effekt nicht da. Ohne die aufwändige Herstellung des Steuerungseffektes bliebe die Veränderung der Umgebungen also ein Stochern im Nebel und der Nudge ein zweigliedriges formatives Objekt. Erst durch die Evidenz wird er zur Triade und erst dann lässt sich eine weitere Geschichte darüber erzählen, wie durch kleine Veränderungen große Wirkungen erzeugt wurden. Und erst durch die Evidenz werden die Aufkleber auf dem Boden zu einem Akteur, der »eine gegebene Situation verändert, indem [er] einen Unterschied macht« (Latour 2007: 123).

## Von Situationen zu Test-Situationen

Statt die Steuerungseffekte ihrer Interventionen in der veränderten Situation zu suchen, blicken die Praktiker:innen also auf Daten, die sie oft erst selbst erheben müssen. Diese Veränderung ähnelt einer Transformation, die Karin Knorr-Cetina Ende der 1980er Jahre mit Blick auf die Astronomie beschrieb. Diese wissenschaftliche Disziplin durchlief damals die Entwicklung von einer Feldwissenschaft zu einer Laborwissenschaft. Durch neue technologische Möglichkeiten der Bildgebung wurde es möglich, Astronomie zu betreiben, ohne je in den Nachthimmel zu schauen. Sie wurde »von der Beobachtung ihres ›Feldes‹ durch Astronomen [...] vollständig unabhängig« (Knorr-Cetina 1988: 88). Durch diese Loslösung der Untersuchungsobjekte aus ihrer »natürlichen Umwelt« (ebd.) wurden sie »im Handlungskontext des Labors *kontinuierlich präsent* gehalten« (ebd.). Gerade weil der Himmel fortan nicht mehr direkt beobachtet wurde, konnten neue Entdeckungen gemacht werden.

Nur dann, wenn Situationen über den Umweg von Daten betrachtet werden, kann mithilfe von Nudges auf sie eingewirkt werden. Die Daten müssen jedoch nicht jedes Mal so mühsam erhoben werden, wie es im Treppenbeispiel der Fall war. Oft liegen sie bereits ›irgendwo‹ vor, und die Herausforderung besteht dann darin, Zugriff auf sie zu erlangen. Denn selbst wenn Daten durch die »Dauer-Sensorisierung« (Nassehi 2019: 80) der Gesellschaft schon existieren, ist nicht ausgemacht, dass sie auch für die Herstellung von Nudges zur Verfügung stehen. Nur wer Zugriff auf die Daten hat, kann sie auch zum Nudging nutzen.

Die Praktiker:innen, deren Arbeit ich untersuchte, operierten als Beratungsunternehmen und mussten folglich bei vielen ihrer Projekte neue Kooperationen etablieren und auch den Zugriff auf die notwendigen Daten über Organisationsgrenzen hinweg verhandeln. Je nachdem, was als relevante Umgebung eines Problems identifiziert wurde, knüpften sie neue Kontakte. Und selbst wenn auf diese Weise Kooperationen zustande gekommen waren, setzte der Zugriff auf Daten weitere Verhandlungsarbeit voraus. Um die Effekte von Interventionen zu ermitteln, die in einem Supermarkt den Absatz nachhaltigen Obstes erhöhen sollten, machte sich eine Mitarbeiterin eines Tages auf den Weg in die Zentrale des Supermarktes. Dort wollte sie die benötigten Daten besorgen:

Ich wende mich Maria zu, die gerade ihre Tasche packt. »So you are hunting for data today, is that right?.« Sie: »Yeah, let's see how that goes! We set up

an appointment for only one hour and I don't think that will be enough.« Ich: »Do you know how they work, what sort of software they use and so on?« Sie: »I have no idea! But I think it will be very complicated to get the data.« Ich: »Do you think you will have to go through the data manually with that person?« Mia: »I need the raw data. And I don't think that they will be willing to give it to me. It is very valuable to them and of course they don't just want to hand it out.« (Notiz 18.09.19)

Bevor sich die Mitarbeiterin mit der für Daten verantwortlichen Person des Supermarktes trifft, zeigt sie sich skeptisch. Sie erwartet Schwierigkeiten und stellt sich auf Überzeugungsarbeit ein, um an die Daten zu kommen. Denn obwohl sie vorliegen und das im Supermarkt stattfindende Kaufgeschehen schon datenförmig verdoppelt ist, bleibt es zu diesem Zeitpunkt unklar, ob auch sie Zugang zu den Daten erhält. An ihrem Problem wird deutlich, dass der Zugriff auf Daten oft ein exklusiver ist, der nicht allen Akteuren mit Steuerungsabsicht offensteht. Sie betont, dass die Daten für die Supermärkte sehr wertvoll seien und sie nicht damit rechne, die Daten bereitwillig ausgehändigt zu bekommen.

Ganz ähnlich stellte sich die Situation im vorigen Kapitel dar, als einer Mitarbeiterin der beobachtende Zugriff auf solche Engstellen verwehrt wurde, die von einem Kreuzfahrtunternehmen kontrolliert wurden. Das Unternehmen hatte eigene Erwartungen an das Verhalten seiner Kund:innen und wollte nicht durch die Interessen Dritter gestört werden. Während die Kooperation mit den Kreuzfahrtunternehmen scheiterte, war sie mit den Supermärkten zustande gekommen. Und selbst der Zugriff auf die Daten gelang meiner Informantin, wie sie mir einen Tag später berichtete:

Maria steht am Waschbecken und spült eine Tasse ab. Ich sage: »So, how did it go yesterday with the guy from the supermarket?« Maria: »Good. It went well, actually. We will get 1.2 million observations on Friday. That's really good.« Ich: »And was it hard to convince him?« Maria: »No, he really did not seem to care. He was just like ›what do you need?‹ and I said ›this and this and this‹ and he gave it to me. I think it was really good, that he was thrown into that task very last minute. We get all purchases of twelve stores for two months, so that's nice!« (Notiz 19.09.19)

Entgegen ihrer eher pessimistischen Erwartungen war die Mitarbeiterin an die Daten gelangt, die sie zum Ermitteln der Wirksamkeit ihrer Interventionen brauchte. Wenn die Intervention Verhaltenseffekte bewirkt hatten, so waren

diese irgendwo in den 1.2 Millionen Beobachtungen zu finden, auf die sie jetzt zugreifen und die sie analysieren konnte. Rückblickend und für einen Zeitraum von zwei Monaten waren die Supermärkte also ihre Test-Situationen geworden. Nicht nur, weil es gelungen war, im Supermarkt bestimmte Interventionen zu implementieren, sondern auch, weil ihr der Zugriff auf die Daten geglückt war, in denen sie nach den Effekten suchen konnte.

Im Treppen- und im Supermarktbeispiel sind zwei Variationen der Herstellung von Evidenz erkennbar. Im ersten Fall muss das beeinflusste Geschehen erst verdoppelt werden, damit der Steuerungseffekt ermittelt werden kann. Im zweiten Fall besteht die Herausforderung darin, Zugriff auf schon bestehende Daten, auf schon bestehende Verdoppelungen, zu erlangen. Beide Beispiele haben indes gemein, dass in ihnen Nudging als eine Bewegung von der datenförmig verdoppelten in die reale Welt und wieder zurück erkennbar wird: Die Bezugsprobleme zeigen sich in Statistiken, wodurch Eingriffe in die Welt plausibel werden, die ihrerseits einen Steuerungseffekt bewirken, der sich aber nur im Bereich der datenförmigen Verdoppelung zeigt. Nudging lässt sich demnach als Technik experimenteller Verhaltensbeeinflussung bezeichnen, die immer von Daten ausgehend, über die Welt, zurück in Daten interveniert.

### V.3 Wiederholen

Im bisherigen Verlauf dieses Kapitels habe ich Nudging als ein Regieren durch Verhaltensexperimente beschrieben, bei dem bestimmte Weltausschnitte laboratisiert werden, damit experimentell in sie interveniert werden kann. Derart betrachtet, ist die Herstellung eines Nudges ein Prozess, bei dem sich die ko-konstitutiven Karrieren von Problem und Lösung als epistemische und technische Objekte zu einem Experimentalsystem fügen, das datenförmige Evidenz generieren kann. Aus einer praxeologischen Innenperspektive auf die Herstellung von Nudges, und geleitet von der Frage, wie die Komponenten des Nudges als formativem Objekt zueinander in Passung gebracht werden, konnte ich zeigen, dass sich Problem und Lösung ko-konstituieren, während die Evidenz als Attraktor auf beide wirkt. Damit ist die Gestalt, die ein fertiger Nudge annimmt, in entscheidendem Maße von der Notwendigkeit bestimmt, Evidenz über die Wirksamkeit dieser sanften Form der Verhaltensbeeinflussung zu erzeugen. Die Art, wie über Nudges berichtet wird, lässt sich also nicht trennen von der Art, wie durch Nudges Verhalten beeinflusst wird. Dies

ist die Pointe, auf die ich bereits am Beginn dieses Kapitels mit Garfinkel verwiesen habe, der auf den integralen Zusammenhang von Berichtsverfahren und der Phänomene, über die sie berichten, hinweist (vgl. Garfinkel 1967: 186ff.).

Mit dem Zusammenwirken von Ko-Konstitution und Attraktion bin ich bei einer Antwort auf die Frage angekommen, wie die drei Komponenten eines Nudges als formativem Objekt zueinander in Passung geraten, und wie es gelingt, einen ›fertigen‹ Nudge herzustellen. Auf den nächsten Seiten möchte ich nun zwei Entwicklungslinien skizzieren, die sich vor dem Hintergrund der These des Nudging als einem Regieren durch Verhaltensexperimente abzeichnen. Hierbei greife ich erneut auf Überlegungen Rheinbergers zurück, der ein Vokabular zur Beschreibung dieser Tendenzen bereitstellt. Im ersten Schritt widme ich mich der Frage der Wiederholbarkeit, bzw. Wiederholungs-Bedürftigkeit von Experimenten und knüpfe daran Überlegungen zur Zukunft des Nudging und der Ausweitung seiner Möglichkeitsbedingungen. Wie kommt man von Einzelexperimenten zu einem Modus des Experimentierens? Im zweiten Schritt thematisiere ich die Tendenz des nachlassenden Evidenzdrucks, wenn Nudges nicht immer wieder erneut aufwändig getestet werden. Diese Tendenz beschreibe ich mit Rheinberger als Übergang vom epistemischen Objekt zur Technologie.

## Vom Experiment zum Experimentieren

Die Praxis der Herstellung von Nudges, die ich als Feldforscher kennenlernen konnte, war eine spezifisch organisierte und spezifisch finanzierte Praxis. Als Beratungsunternehmen stellten meine Informant:innen Nudges für Auftraggeber:innen her, die mit Problemen an sie herantraten und oft zum ersten Mal auf Nudging als Werkzeug zurückgriffen. Mir zeigte sich die Herstellung von Nudges also als projektförmig organisiert, und immer wieder mussten meine Informant:innen die Bedingungen erst schaffen, unter denen Nudging möglich ist. Immer wieder mussten sie – je nach Problem – neue Umgebungen so unter ihre Kontrolle bringen, dass sie veränderbar sind und Messungen ermöglichen. Damit glich jedes Projekt einem einzelnen Experiment, bei dem sehr viel Organisations- und Abstimmungsarbeit nötig war, bis alles wie benötigt arrangiert war. Wenn ich diese Herstellungspraxis nun mit Überlegungen zu Experimentalsystemen konfrontiere, fällt Rheinbergers Betonung auf, dass in der Praxis der Wissenschaften Experimente eigentlich nie isoliert oder als Einzelexperimente auftauchen. Er spricht sich sehr deutlich gegen »die üb-

liche Auffassung vom Experiment als Testverfahren für Hypothesen« (Rheinberger 2006: 24) aus, die der von mir beobachteten Herstellungspraxis doch zugrunde zu liegen scheint.

Das ist seltsam. Denn die Akribie, mit der meine Informant:innen die Steuerungseffekte ihrer Interventionen sichtbar machen; die Mühsal, mit der sie teils händisch Daten erheben, um Verhaltenseffekte veränderter Umgebungen zu zeigen, gleicht doch einer Sisyphusarbeit, wenn Rheinberger im Anschluss an Ludwik Fleck betont, »wie wenig ein Einzelexperiment beweist und zwingt« (Fleck 1980: 126 zit. nach Rheinberger 2006: 24). Statt mit Einzelexperimenten habe ein Forscher es in der Praxis stets »mit einer Experimentalanordnung [zu tun], die er so entworfen hat, daß sie ihm Wissen zu produzieren erlaubt, das er noch nicht hat« (Rheinberger 2006: 24). Ein so verstandenes Experimentieren entspricht weniger einem isolierten Testen klar formulierter Hypothesen, als einem ›Herauspräparieren‹ oder ›Hervorlocken‹ von Untersuchungsgegenständen (vgl. Haus 2021: 2), das in Form von Einzelexperimenten undenkbar ist. Auch die anderen Theoretiker:innen, die in diesem Kapitel bisher zu Wort kamen, betonten die Serialität von Experimenten. Laborisierung ist Callon et al. zufolge beispielsweise ein »interminable undertaking, always starting up again« (Callon/Lascoumes/Barthe 2009: 67). Auf ähnliche Weise heben Marres und Stark mit Blick auf Tests hervor, dass es nie bei *einem* Test bleibe: »Tests are generative, they stimulate further testing.« (Marres/Stark 2020: 430)

Experimente setzen also Experimente voraus und weitere Experimente in Gang. Ich möchte dies anhand einer Interviewsequenz verdeutlichen, in der mir mein Gesprächspartner von einem Projekt berichtete, bei dem er mithilfe von Nudges den Absatz nachhaltiger Gerichte in einer Kantine erhöhen wollte. Seiner Vermutung nach sollte sich ein besonders großer Effekt dann einstellen, wenn das entsprechende Gericht an der ›besten Position‹ ausgegeben wurde. Doch in den Daten konnte er den vermuteten Effekt nicht finden:

B: Ich hab mir gedacht, dass die Ausgabeposition den größten Einfluss an der ganzen Geschichte hat. Dass die Zahlen jetzt nicht so wirklich aussagekräftig sind, kann natürlich auch daran liegen, dass wir nur eine Nullmessung gemacht haben und diese Nullmessung ein bisschen verzerrt ist. Weiß ich natürlich nicht. Beim nächsten Mal würde ich dann vorschlagen, dass wir ungefähr zwei oder drei Nullmessungen machen und dann den Durchschnitt davon bilden, gibt vielleicht ein reineres Ergebnis, das heißt vielleicht sind ja auch alle effektiv und nur die Nullmessung war verzogen, also das wär

ein Traum, aber kann man im Endeffekt nicht so richtig abschätzen, müsste man nochmal untersuchen, aber an sich bin ich schon der Überzeugung, dass wenn so ein nachhaltiges Gericht auf der besten Position angeboten wird, dass es dann auch mehr abgesetzt wird. (Interview 01.02.2018)

In milder Form deutet sich hier an, was Harry Collins (1981) als »experimenter's regress« beschrieben hat: Wie lässt sich herausfinden, dass die Ausgabeposition den größten Effekt hat? Man muss ein Experiment durchführen, das diesen Effekt misst. Wie lässt sich feststellen, dass das Experiment korrekt durchgeführt wurde? Das hängt davon ab, ob es das richtige Ergebnis ausspuckt. Wie wird klar, dass dies das richtige Ergebnis ist? Dazu müsste ein weiteres Experiment durchgeführt werden. Und so weiter...

Usually, successful practice of an experimental skill is evident in a successful outcome to an experiment, but where detection of a novel phenomenon is in question, it is not clear what should count as a »successful outcome« – detection or non-detection of the phenomenon. Thus arguments concerning the existence of the phenomenon turn, not upon experimental results, but upon what comes to count as a »well-done experiment«. (Collins 1981: 34)

Am experimenter's regress wird deutlich, dass sich aus einem Einzelexperiment wenig lernen lässt, und dass Experimentieren aus sich heraus kein Ende kennt. Vor dem Hintergrund dieser Überlegungen wirkt es sehr merkwürdig, wenn die aufwändig erarbeiteten Experimentalanordnungen, die im Zuge der Herstellung von Nudges entstehen, wieder verworfen werden, sobald Evidenz hergestellt und der Nudge als dreigliedriges formatives Objekt komplettiert ist. Dann endet das Projekt, ein Bericht mit Balkendiagrammen kann geschrieben, und eine Rechnung an die Auftraggeberin ausgestellt werden. Aber ist damit das Potenzial eines Regierens durch Verhaltensexperimente ausgeschöpft? Die Herstellung von Einzelexperimenten, die ich beobachten konnte, ist eine sehr spezifische Form, Nudging zu betreiben. Aus der Perspektive meiner Informant:innen, die Nudges verkaufen und sich dabei auf ihr Image als *wissenschaftliche* Beratungsgesellschaft stützen, ist es völlig plausibel eine verbreitete Vorstellung von Wissenschaftlichkeit zu bedienen, die »der Auffassung vom Experiment als Schiedsrichter verhaftet [ist]« (Rheinberger 2006: 24).

Doch wenn wir nun vom dem konkreten Herstellungskontext abstrahieren, in dem ich Nudging als teilnehmender Beobachter kennenlernen konnte, wie könnte dann eine Zukunft des Nudging aussehen, in welcher der ange-

strebte Machteffekt des Nudging nicht durch die immer neue Schaffung von Experimentalbedingungen gebremst wird? Vor dem Hintergrund von Rheinbergers Überlegungen liegt es nahe zu argumentieren, dass Nudging vor allem dort sein Versprechen des wirksamen Regierens einlösen kann, wo stabile Experimentalbedingungen herrschen. Also dort, wo ohne viel Aufwand Umgebungen verändert werden können, und wo etablierte Verfahren der Überwachung herrschen, die eine datenförmige Verdoppelung der Welt gewährleisten.

Erste Anzeichen sich stabilisierender Experimentalbedingungen konnte ich als teilnehmender Beobachter immer dann erkennen, wenn ich eine Informantin an eine Behörde begleitete, wo sie seit mehreren Jahren als Beraterin tätig war und im Laufe der Zeit eine Vielzahl von Projekten durchgeführt hatte. Eines Vormittags kamen wir nach einem langen und konzentrierten Projekttreffen miteinander ins Gespräch, und sie begann darüber zu reflektieren, wie sich die Arbeit hier in der Behörde von ihrem sonstigen Projektalltag in der Agentur Agentur unterschied.

B: Here [in der Behörde, T.S.] I am doing (..) what I'm good at. And I'm only doing the things where I (...) where I'm most productive or like where I contribute the most. And I'm not the one sitting and calling a lot of people and asking about data. I prefer participating in meetings, writing up experiments, experimental guidelines and protocols. Or writing up solutions. And all the other things around it I don't have to deal with. So yeah, I like that better. And it makes more sense. (Notiz 14.10.19)

Während im Agenturalltag ein Großteil ihrer Arbeit daraus bestand, Projektpartner:innen hinterherzutelefonieren und den Zugriff auf Daten zu verhandeln, also Experimentalbedingungen zu schaffen, konnte sich meine Informantin in der Behörde auf die Durchführung von Experimenten konzentrieren. Hier konnte sie die Aspekte voraussetzen, die sie im Projektgeschäft der Agentur immer wieder aufs Neue erarbeiten musste.

Dieser routinisierte Modus des Experimentierens war nur möglich, weil in dieser Behörde bereits eine Vielzahl erfolgreicher Nudging-Projekte durchgeführt worden war, und die Verantwortlichen in der Organisation nach und nach davon überzeugt werden konnten, dass Nudging ein hilfreiches Werkzeug für ihre Probleme ist:

I think they're slowly getting convinced here. But it has taken a long time. And we've made a lot of experiments and had a lot of good results. And now they organized it differently. So now they have a team. And I don't know if it is going to last forever. But they're trying to get like a really good team that works across the organization. So we've spend eight years or so doing experiments. And by now like I think the organization and also the people on the top are convinced that it makes sense. Because we've been able to create results.

It has only been possible because we can produce the results. And it's not everything that works out, like sometimes we don't get a good result. But continuously working with data and doing experiments and having these hard numbers saying we actually make a difference. (Notiz 14.10.19)

Die schrittweise Durchsetzung des Nudging, die meine Informantin hier vor dem Hintergrund ihrer Erfahrungen in der Behörde beschreibt, macht für mich nun auch die Agenturarbeit der Durchführung von Einzelexperimenten besser nachvollziehbar. Die aufwändig durchgeführten Projekte gleichen Pilotstudien und Prototypen, mit denen Nudging als Innovation der Verhaltensbeeinflussung mit immer neuen Auftraggeber:innen ausprobiert wird. Sie können sich dann für Folgeprojekte entscheiden, und sich – wie die Behörde – Projekt für Projekt das Experimentieren als einen etablierten Modus der Verhaltensbeeinflussung aneignen.

Wie meine Informantin betont, gelingt diese sukzessive Etablierung des Nudgings jedoch nur, wenn kontinuierlich gute und überzeugende Ergebnisse zeigen, »[that] we actually make a difference« (Notiz 14.10.19). Somit wird deutlich, dass die Herstellung von Evidenz nicht nur Notwendigkeit einer sanften Form der Verhaltensbeeinflussung, sondern dass Evidenz auch ein maßgeblicher Faktor in der Durchsetzung des Nudging ist. Thomas Scheffer betont, dass strategische Überlegungen der Durchsetzung einer Herstellungspraxis nicht äußerlich sind, sondern auch in der Praxis selbst mitbearbeitet werden. Etwa bei »Anwält:innen [sic], die nicht nur eine ordentliche Fallarbeit, sondern auch Rechtsentwicklung betreiben oder Politiker:innen, die nicht nur annehmbare Positionen liefern, sondern diese auch gegen Widerstände verfechten« (Scheffer 2020: 222). Demnach ist die Arbeit an der Evidenz zentraler Motor einer Durchsetzung des Nudging und eines Übergangs vom Einzelexperiment zum etablierten Modus experimenteller Verhaltensbeeinflussung.

Der hier skizzierte Prozess einer sukzessiven Durchsetzung des Nudging, in dessen Verlauf eine Organisation entsprechende Stellen schafft und Projekt

für Projekt in einen Modus des Experimentierens kommt, gleicht einer Bewegung in Richtung seines »angepeilten Machteffekt[s]« (Scheffer 2020: 225). Wenn innerhalb einer Organisation etablierte Arbeitsbeziehungen herrschen und somit der Zugriff auf Daten und die Veränderbarkeit von Prozessen und Umgebungen gewährleistet ist, dann sinkt der Aufwand der Herstellung von Nudges im Vergleich zum Agenturgeschäft. Damit gelingt womöglich der Übergang vom Einzelexperiment zum seriellen Experimentieren. Diesen Prozess weiter nachzuverfolgen, wäre Stoff für ein weiteres Forschungsprojekt, für das mir an dieser Stelle die Ressourcen fehlen.

Statt der Verbreitung und Etablierung des Nudging also weiter nachzugehen, möchte ich einen theoretischen Ausblick in Richtung eines Horizontes werfen, in dem der angepeilte Machteffekt dieser Regierungstechnologie weitestgehend verwirklicht sein könnte. Worauf läuft die Durchsetzung des Nudging hinaus? Im Verlauf dieser Studie ist die Affinität des Nudging zum Digitalen immer wieder aufgeflackert. Und in diesem Kapitel habe ich mit Blick auf digitale Technologien bereits zwei potenzielle Endpunkte skizziert, in denen Nudging sein Potenzial voll entfalten könnte. Marres und Stark (2020) sprachen von »total test environments« (436), auf die sich Gesellschaften hinbewegen:

[T]he operations that produce today's total test environment consist of minor modifications in the environments in society so as to render *the setting capable of data capture, analysis, and feed-back* – that is, to equip it as a test environment, to enable representation *and* intervention – even if aspirationally – on a more or less durable basis. (Marres/Stark 2020: 436)

Ähnlich argumentierte Mühlhoff, dass sich »[d]urch den fast flächendeckenden Einsatz von Tracking-Techniken [...] das Internet [...] zu einem behaviorellen Echtzeitlabor [entwickle]« (Mühlhoff 2018: 562). Bei beiden dieser Diagnosen fällt auf, dass die Autor:innen mit den »total test environments« und dem »behaviorellen Echtzeitlabor« jeweils dystopische Szenarien entwerfen, die zwar noch nicht verwirklicht seien, in deren Richtung aktuelle gesellschaftliche Entwicklungen jedoch deuteten. Das, was allgemein als Digitalisierung verstanden wird, scheint Bedingungen mit sich zu bringen, unter denen Nudging prosperieren kann.

Wenn ich vor dem Hintergrund meiner Forschungserkenntnisse nun mit Nassehi danach frage, »[f]ür welches Problem die Digitalisierung eine Lösung ist« (Nassehi 2019: 17), so ist die Antwort klar: Sie ermöglicht meinen Infor-

mant:innen perspektivisch, Nudging mit viel weniger Aufwand zu betreiben und sie lässt eine Gesellschaft in Reichweite rücken, in welcher eine konsequent experimentelle Verhaltensbeeinflussung erst möglich wird. Zum einen, weil – wie in den obigen Diagnosen angedeutet – der großflächige Einsatz von Trackingtechnologien die datenförmige Verdoppelung der Welt quasi automatisiert, die nötig ist, um die latenten Steuerungseffekte von Nudges zu erkennen. Zum anderen aber auch – und dies lässt sich vor allem vor dem Hintergrund der praktischen Schwierigkeiten des Nudging nachvollziehen –, weil sich das Digitale durch eine spezifische Ontologie (vgl. Kallinikos/Aaltonen/Marton 2013) auszeichnet, in der Veränderungen von Umgebungen weitaus weniger aufwändig sind, als in der analogen Welt.

### Vom epistemischen Objekt zur Technologie

Neben der Ausweitung und Stabilisierung von Experimentalbedingungen, durch die es gelingt, Nudging mit weniger Aufwand zu betreiben, konnte ich während meiner Feldforschung noch eine andere Tendenz feststellen. Denn immer lernte ich auch Beispiele kennen, in denen keine Experimente durchgeführt und keine Evidenz generiert wurde. Dies kann verschiedene Gründe haben. In folgendem Beispiel hatte eine Informantin beispielsweise behavioral informierte Gestaltungsempfehlungen abgegeben, die im konkreten Fall jedoch nicht evaluiert wurden, weil das Thema politisch zu heikel war:

B: Also bei dem Projekt haben wir kein Trial gemacht am Ende. Also es wurde nichts evaluiert wir haben quasi eine Liste von verschiedenen Empfehlungen abgegeben was man machen könnte. Also ein Report geschrieben genau und das dann dem Kunden überliefert und der Kunde hat dann entschieden, dass das politisch zu heikel ist das zu publizieren und deswegen wurde das auch nicht mehr weiterverfolgt. (Interview 14.03.19)

In anderen Fällen schien der Evidenzdruck aber auch abzunehmen, wenn sich Nudging innerhalb von Organisationen stärker etabliert hatte, und ein generelles Vertrauen in die Wirksamkeit des Nudging bestand. Dann mussten die Praktiker:innen nicht jeden Nudge erneut durch Experimente untermauern und konnten stattdessen auf andere Quellen von Evidenz zurückgreifen. Bei einem Fachaustausch, an dem ich während meiner Feldforschung teilnehmen konnte, berichteten zwei Nudging-Beauftragte eines europäischen

Flughafens davon, dass sie ihre Interventionen kaum noch durch Experimente evaluierten:

Eine Mitarbeiterin sagt, dass sie kaum mehr ›Follow-Ups‹ machen würden, weil ihnen dazu schlicht die Zeit fehle. Ihrer Vorgesetzten vertrauten ihnen, und meist würde es reichen, wenn die Beschwerden bezüglich der Probleme, denen sie sich widmen, nachlassen würden. Beispielsweise habe es das Problem gegeben, dass oft Personen ohne Boarding Pass in Geschäften gelandet seien, in denen sie ohne Pass nicht einkaufen dürften. Um die implementierte Intervention zu validieren, habe es dann ausgereicht, dass die Beschwerden der Händler:innen aufgehört hätten. (Notiz 24.09.19)

In diesem Beispiel schien die gemessene Evidenz insofern obsolet zu werden, als dass anekdotische Evidenz ausreichte, um die Lösung eines Problems zu bestätigen. Wenn nach der Implementierung einer Lösungsidee die Problembeschwerden nachlassen, schien dies zu genügen. Vertrauen in die Wirksamkeit von Seiten der Vorgesetzten, und etablierte Kommunikationswege innerhalb der Organisation, entlang derer andere Formen von Evidenz beschafft werden konnten, deuten hier auf eine viel stärkere Integriertheit des Nudging hin. Wo Vertrauen herrscht, scheint die Evidenz weniger dringlich zu sein.

Solche Beispiele des nachlassenden Evidenzdrucks lassen sich mit Rheinberger als Übergang von epistemischen Dingen zur Technologie denken, wenn »ausreichend stabilisierte epistemische Dinge [...] als technische Bausteine in eine bestehende Experimentalanordnung eingefügt« (Rheinberger 2006: 29) und dort »zu konstituierenden Momenten der experimentellen Anordnung [werden]« (Rheinberger 1992: 70). Wenn die Wirkung bestimmter Nudges oft genug nachgewiesen und stabilisiert worden ist, dann werden die entsprechenden Interventionen auch woanders und ohne experimentellen Aufwand implementiert, weil klargeworden ist, dass damit bestimmte Effekte erzielt werden können. Ein Informant schilderte mir dies wie folgt:

Er kenne es von vielen Nudge Units so, dass sie einfach eine Studie heranzögen, bei der beispielsweise der *present bias* eine Rolle spiele, und die Intervention aus dieser Studie dann auf einen anderen Kontext anwendeten. Das werde dann häufig überhaupt nicht getestet, sondern einfach implementiert. (Notiz 15.12.2019)

An diesem Beispiel ist der Verweis auf frühere Studien bemerkenswert. Der Evidenzdruck scheint dort nachzulassen, wo sie durch den Verweis auf bereits

ermittelte Evidenz substituiert werden kann. Auf diese Weise scheint sich ein nudgingspezifisches Regierungswissen herauszubilden, das Gestaltungsprinzipien übernimmt, aber nicht mehr den aufwändigen Prozess der Herstellung von Nudges durchlaufen muss, den ich in dieser Studie nachgezeichnet habe. Dies ist die andere Tendenz der Verbreitung des Nudging. Statt einer Stabilisierung von Experimentalbedingungen, wodurch ein Modus des Experimentierens möglich wird, besteht sie in einer Verbreitung erprobter und als Verhaltenstechnologien stabilisierter Gestaltungsprinzipien. Rheinberger betont, dass die zur Technologie geronnenen epistemischen Objekte dann damit beginnen, »selbst die Spannweite der Fragen zu bestimmen, die innerhalb des Systems gestellt werden können« (Rheinberger 1992: 70f.) – *if you have a hammer, everything looks like a nail.*

#### V.4 Zwischenfazit zur Arbeit an der Evidenz

Damit bin ich am Ende der Rekonstruktion der Arbeit an der Evidenz angekommen. Ausgehend von der Beobachtung, dass die Sanftheit ihrer Regierungsbemühungen für die Praktiker:innen selbst zum epistemischen Problem wird, rekonstruierte ich die aufwändigen Verfahren, mit denen sie die Steuerungseffekte ihrer Nudges sichtbar machen. Ich deutete Nudging als ein Regieren durch Verhaltensexperimente, das bestimmte Weltausschnitte laborisiert, um direkt experimentell in sie intervenieren zu können. Statt *in* der Welt Experimente durchzuführen, besteht Nudging also aus einem Experimentieren *an* der Welt. Mit Blick auf Nudges als formative Objekte konnte ich so zeigen, dass die Herstellung eines Nudges aus der Ko-Konstitution von Problem und Lösung besteht, die sich als epistemische und technische Dinge zueinander fügen, während mit der Evidenz als Attraktor eine zusätzliche Kraft auf sie wirkt, die den Verlauf ihrer Objektkarrieren entscheidend mitbeeinflusst.

Unter Rückgriff auf das Theorem der datenförmigen Verdoppelung der Welt arbeitete ich zudem heraus, dass Nudging darauf angewiesen ist, seine latenten Steuerungseffekte als Muster in Daten zu suchen. Denn weil es sich bei Nudging lediglich um die Beeinflussung von *Verhaltenswahrscheinlichkeiten* handelt, sind die Effekte in den beeinflussten Situationen nicht direkt beobachtbar. Stattdessen können sie nur zeitverzögert, nämlich wenn die Daten vorliegen, und andernorts, nämlich mit Blick auf Daten ermittelt werden. Damit wurde auch klar, dass das Potenzial von Nudging vor

allem dort liegt, wo durch automatisierte Datenerhebungs-, Sammlungs- und Auswertungstechnologien die Bedingungen gegeben sind, um Experimente ohne viel Aufwand zu wiederholen und in einen Modus der experimentellen Verhaltenssteuerung umzuschalten. Neben dieser Ausweitung der Möglichkeitsbedingungen des Nudging zeigte sich auch eine andere Tendenz des nachlassenden Evidenzdrucks. Wenn die Wirksamkeit bestimmter behavioraler Gestaltungsprinzipien oft genug bewiesen worden war, verbreiteten sich diese mit Verweis auf frühere Ergebnisse, ohne erneut getestet zu werden. Sie wurden also von epistemischen Dingen zur Technologie, was mit einer Verfestigung der lösbaren Probleme einhergeht: »[B]is auf weiteres [gibt es] nur solche Probleme, die sich mit Betriebsmitteln auch bearbeiten lassen« (Scheffer 2019: 345).

Verlernen wir also im Zuge der Durchsetzung des Nudging, Probleme anders als als Verhaltensprobleme zu betrachten? Schrumpft dabei die Spannweite der Fragen, die wir überhaupt zu stellen in der Lage sind, auf die Suche nach problematischem Verhalten und Ursachen in der Umgebung zusammen? Es würde dem Charakter dieser Studie nicht gerecht werden, hier nun mit einer alarmistischen These zu enden, die dem Nudging grundsätzlich seine Berechtigung abspricht. Ich halte es eher für angemessen, darüber nachzudenken, wofür das Konzept geeignet sein könnte. Mit anderen Worten: Es geht mir darum, »jenseits einer generellen Apparate- bzw. Machtkritik, [...] die Bearbeitungskapazitäten in ihren je aktuellen Bedingtheiten zu erfassen« (Scheffer 2020: 233).