Die Grammatik des Vertrauens: Eine Untersuchung in interpersoneller und epistemischer Hinsicht

Inaugural-Dissertation

zur Erlangung des Doktorgrades der Philosophischen Fakultät der Universität zu Köln

im Fach Philosophie

vorgelegt von

Jörn Wiengarn

geb. am 20.05.1990

in Stadtlohn

Köln, den 13.10.2021

Diese Dissertation wurde Januar 2022 von der Philosophischen Fakultät der Universität zu Köln angenommen.

Inhaltsverzeichnis

I. Erster einleitender Teil.....4

1. Einleitung.....6
1.1 Thema und Fragestellung: Vertrauen in Personen.....6
1.2 Methodische Bemerkungen: Orientierung am gewöhnlichen Begriffsverständnis.....14
1.3 Das Vorgehen im Einzelnen.....21

2. Der generische Begriff des Vertrauens: Eine erste Annäherung.....24
2.1 Eine formale Definition: Vertrauensakt und vertrauende Haltung.....24
2.1.1 Vertrauen und Risikobewusstsein.....26
2.1.2 Vertrauen als zuversichtliche Haltung.....30
2.2 Interpersonelles Vertrauen als Vertrauen *sui generis*.....31

II. Teil: Zur interpersonellen Grammatik von Vertrauen.....33

3. Die Rational Choice-Theorie des Vertrauens.....34
3.1. Ziel des Kapitels.....34
3.2 Eine kurze Einführung in die Rational Choice- und Spieltheorie des Vertrauens.....35
3.3 Exkurs: Der humeanische Kern der Rational Choice-Theorie.....38
3.4 Kritik des Vertrauensbegriffes der Rational Choice-Theorie.....42
3.4.1 Die normative Erwartung des Vertrauenden.....43
3.4.2 Die Einstellung des Vertrauensnehmers.....47
3.5 Zwischenfazit: Interpersonelles Vertrauen als Vertrauen *sui generis*.....50

4. Die Axiologie interpersonellen Vertrauens.....52
4.1 Der Wert in Vertrauen und Vertrauenswürdigkeit: Ein Sortierungsversuch.....54
4.2 Normativität, Handlungsgründe und der Wert von Vertrauenswürdigkeit.....64
4.3 Exkurs: Von der humeanischen zur wertbezogenen Rationalitätskonzeption.....65
4.4 Die teleologische Werttheorie: Eine instruktive Gegenposition.....68
4.5 Zwei Einwände gegen die teleologische Lesart.....70

5. Die relationale Normativität zwischen den Vertrauenspartnern.....74
5.1 Relationale Normativität.....75
5.2 Der normative Kontext von Vertrauen.....79
5.3 Vertrauenswürdigkeit: Eine tugendethische Konzeption.....89
5.3.1 Absichtsbezogene Konzeptionen von Vertrauenswürdigkeit.....90
5.3.2 Vertrauenswürdigkeit und normative Anforderungen.....94
5.3.3 Vertrauenswürdigkeit als Tugend.....99

5.3.4 Vertrauenswürdigkeit und Kompetenzen....102

6. Vertrauen als Anerkennung....103
6.1 Das Problem der Sozialkalkulation....106
6.2 Anerkennung und Zeugenschaft: Eine Kritik an Richard Moran....113
6.3 Vertrauen als Anerkennung....123
6.3.1 Die Anerkennungsthese des Vertrauens....124
6.3.2 Die evaluative Komponente vom Vertrauen....129
6.3.3 Die Abgrenzung zur sozialkalkulierenden Einstellung....132
6.3.4 Diskussion zweier Einwände....133
6.4 Vertrauen und die Beziehung der beiden Vertrauenspartner....136
6.5 Zusätzliche Bemerkungen....140
6.6 Zwischenfazit: Das explikatorische Potenzial der Anerkennungstheorie....142

III. Teil: Zur epistemischen Grammatik von Vertrauen....143

7. Die epistemische Struktur von Vertrauen....144
7.1 Einleitung: Die Relation zwischen Vertrauen und epistemischen Gründen....144
7.2 Der Kognitivismus des Vertrauens....146
7.3 Epistemische Gründe und der epistemische Anspruch im Vertrauen....154
7.4 Vertrauen als festgelegte Haltung....161

8. Vertrauen und Reflexion....163
8.1 Die Emotionstheorie des Vertrauens....165
8.2 Die Tugendtheorie des Vertrauens....169
8.3 Vertrauen und Misstrauen als habitualisierte Praxis....173
8.3.1 Praxis als kognitiv strukturierte Tätigkeitsform....174
8.3.2 Der Automatizitätcharakter von Praktiken....177
8.3.3 Lernprozesse und Verantwortlichkeit: Die Idee kontrafaktischer Sensitivität....178
8.4 Vertrauen als wesentlich habitualisiert....186
8.5 Zwischenfazit....190

IV. Teil: Die Beziehung zwischen interpersoneller und epistemischer Seite....191

9. Integration der epistemischen und der interpersonellen Seite....192
9.1 Die Dichotomie zweier Sichtweisen auf Vertrauen....192
9.1.1 Der additive Lösungsansatz....197
9.1.2 Exkurs: Zum Lösungspotenzial von *Entitlement*-Ansätzen....202
9.1.3 Der innere Zusammenhang von Anerkennung und epistemischen Gründen....209
9.2 Epistemische Permissivität....217

10. Schluss: Vertrauen an der Schnittstelle zwischen Ethik und Epistemologie....221

Bibliographie....223

Lebenslauf des Autors....239

Die vorliegende Promotion wäre ohne die großzügige Unterstützung, intellektuelle Anregung und zusprechende Ermutigung durch viele Personen nicht möglich gewesen. Ihnen möchte ich an dieser Stelle meinen Dank aussprechen. Dabei danke ich in erster Linie ganz herzlich Christine Chwaszcza, die ich als hervorragende Erstbetreuerin erlebt habe. Sie hatte stets ein offenes, interessiertes Ohr, kluge, und auch kritische, aber hilfreiche Ratschläge parat und mich nach allen Kräften und Möglichkeiten bei der Verfolgung meiner intellektuellen Interessen unterstützt. Ich danke auch meinem Betreuer Wilfried Hinsch, mit dem ich äußerst intensive, kluge und erkenntnisreiche Gespräche über die Themen der Arbeit, und damit viele Verbesserungen derselben verdanke. Ein besonderer Dank gilt Andreas Kaminski: Nicht nur, weil ein Seminar bei ihm an der TU Darmstadt am Anfang meines Interesses am Vertrauensthema steht; seine wegweisenden Überlegungen zum Vertrauensthema, die ich auch in zahlreichen persönlichen Gesprächen mit ihm diskutieren konnte, haben maßgeblich Einfluss in dieser Arbeit gefunden. Ich verstehe diese Arbeit als im Geiste der Überlegungen von ihm stehend. Ich danke außerdem ausdrücklich der a.r.t.e.s. Graduate School der Universität zu Köln, und stellvertretend ihrem Direktor Andreas Speer. Ihrem Vertrauen in meine Arbeit habe ich nicht nur eine großzügige finanzielle Unterstützung zu verdanken, sondern auch eine äußerst angenehme und anregende Umgebung zum Schreiben einer Dissertation. Nicht zuletzt haben sich auf diesem Wege viele bedeutende Freundschaften für mich ergeben. Zuletzt möchte ich von Herzen meinen Eltern danken. Sie haben mir unerschütterliche emotionale Unterstützung beim Verfassen dieser Arbeit gegeben. Beide sind zudem nicht nur ausgezeichnete praktische Kenner der Vertrauensgrammatik, meine Mutter ist darüber hinaus auch eine beeindruckende Kennerin der deutschen Grammatik. Ihnen beiden ist diese Arbeit gewidmet.

I. Erster einleitender Teil

1. Einleitung

1.1 Thema und Fragestellung: Vertrauen in Personen

> „The 'riddle' of trust stems from a tendency to analyse it in confusing categories rather than from anything inherently mysterious about it.“[1]

Vertrauen, so wurde schon oft festgestellt, ist nicht nur ein ubiquitäres und fundamentales Phänomen, sondern kennt auch eine Mannigfaltigkeit von Instanzen, auf die es sich beziehen kann. Grundlegend scheint es, als gäbe es keinen Bereich unseres Wissens und Handelns, der nicht in irgendeiner Form Vertrauen zur oft unbemerkten Voraussetzung hat. Um uns in der Welt kognitiv und in unserem Handeln überhaupt orientieren zu können, sind wir irreduzibel darauf angewiesen zu vertrauen. So sehr Vertrauen ein durchdringendes Grundkonstituens sämtlicher menschlicher Lebensbereiche darstellt, so sehr kennt es aber auch eine Multiplizität ganz unterschiedlicher Bezugspunkte: Sei es als ein grundlegendes Vertrauen in unsere Mitmenschen überhaupt, als Vertrauen in Expert:innen und Funktionsträger, als Vertrauen in gesellschaftliche Institutionen, als Vertrauen in Technik oder schlichtweg als Vertrauen in die Regelmäßigkeit bestimmter Abläufe.

Innerhalb dieser Menge an möglichen Bezugsformen von Vertrauen, sticht nun eine ganz grundlegende Unterscheidung hervor, die auch in der philosophischen Diskussion um das Vertrauensphänomen im Zentrum steht. So bezieht sich unser Vertrauen einerseits auf bestimmte Umstände oder Abläufe: Wir haben etwa Vertrauen darauf, dass die Technik funktioniert, dass der Wetterbericht stimmt, oder noch grundlegender, darin, dass der Boden auf dem ich laufe mein Gewicht trägt, oder dass morgen wieder die Sonne aufgehen wird. Andererseits vertrauen wir aber auch *anderen Personen*, und zwar in diversen Rollen und innerhalb verschiedener Kontexte: Ich vertraue etwa meinen Freunden, dem Nachrichtensprecher, wissenschaftlichen Expert:innen[2] oder dem Verkäufer an der Kasse.

Im Zentrum dieser Arbeit steht diese Form des Vertrauens in andere Personen. Dessen Spezifität herauszuarbeiten wird das zentrale Ziel sein. Die zentrale Fragestellung lautet, was es

1 Lagerspetz 1998, 27.

2 In der folgenden Arbeit werde ich mehr oder weniger zufällig zwischen dem grammatisch männlichen und grammatisch weiblichen Geschlecht wechseln, sofern ich mich dabei auf imaginäre oder unbestimmte Personen (etwa Akteur/innen, Vertrauenspartner/innen) beziehe. Diese sind natürlich generisch zu verstehen, stehen also stellvertretend für alle Geschlechter. Die mehrere Geschlechter nennende Formulierung werde ich dann verwenden, wenn ich auf reale Personen referiere (beispielsweise: die Autor:innen).

bedeutet jemandem zu vertrauen. Dabei wird uns die grundlegende Annahme leiten, dass es sich dabei um eine Vertrauensform *sui generis* handelt, das heißt, dass sich Vertrauen in Personen nicht einfach nur durch seinen spezifischen Gegenstandsbereich auszeichnet. Der hier im Fokus stehende Typus personenbezogenen Vertrauens eignet also nicht bloß der triviale Umstand, dass es sich eben auf Personen bezieht. Vielmehr zeichnet es sich durch eine charakteristische interpersonelle *Form* aus, in der sich Vertrauensgeber und Vertrauensnehmer aufeinander beziehen. Wie wir im Laufe dieser Arbeit sehen werden, kennt die bisherige Vertrauensliteratur zwar auch ein dezidiertes Interesse für die personenbezogene oder interpersonelle Vertrauensform, weist aber noch Defizite dabei auf, dessen charakteristische Form zu konzeptionalisieren. Diese besser zu verstehen und klarer zu fassen, und zwar, erstens in seiner anerkennungstheoretisch-interpersonellen und zweitens in seiner epistemischen Dimension, stellt deshalb das zentrale Anliegen dieser Arbeit dar.

Um das Phänomen, um das es uns in dieser Arbeit zu tun ist, klarer zu identifizieren, fangen wir jedoch zunächst mit der Betrachtung eines Beispiels an. Zuweilen sieht es nämlich so aus, als würde sich unser Vertrauen auf andere Personen beziehen, tut dies jedoch nicht in der genuin interpersonellen Art und Weise, die im Fokus dieser Arbeit steht. Mit anderen Worten: zuweilen referiert unser Vertrauen zwar auf eine andere Person, jedoch ohne, dass wir *der anderen Person* vertrauen. Um den Unterschied zu verdeutlichen stelle man sich folgenden Fall vor: Angenommen mein Nachbar Peter sei ein besonders eingefleischtes Gewohnheitstier mit einem rigide durchritualisierten Tagesablauf. Zu seinen täglichen Routinen gehöre unter anderem, dass er jeden Morgen exakt um die gleiche Uhrzeit aufsteht und seine Wohnung jeden morgen pünktlich um acht Uhr verlässt. Stellen wir uns weiterhin vor, er schließe seine Wohnungstür dabei stets mit einem derart lauten Knall, dass ich verlässlich um acht Uhr von ihm geweckt werde. Schon seit Jahren verhalte es sich so, dass er mich Morgen für Morgen um exakt die gleiche Uhrzeit aus dem Bett schrecken lässt. Eines Tages nun trete ein glücklicher Zufall ein: Wegen eines wichtigen Termins muss ich tatsächlich am nächsten Morgen um acht Uhr aufstehen. In diesem Fall mag ich mir sagen: „Wunderbar, ich muss mir nicht einmal einen Wecker stellen – ich kann schließlich darauf vertrauen, dass *Peter* mich pünktlich wecken wird“.

Diese Redeweise scheint in dem beschriebenen Szenario durchaus angemessen zu sein. Die Frage stellt sich jedoch: Würde ich auch sagen können, dass ich *Peter* vertraue, dass er mich morgen wecken wird? Es scheint, dass diese zweite Ausdrucksweise hier nicht ganz treffend wäre oder allenfalls mit einem ironischen Beiklang gesagt werden könne. Ich vertraue zwar darauf, *dass* Peter mich wecken wird, aber ich vertraue nicht (wirklich) *Peter* diesbezüglich. Ich vertraue also

auf *etwas* (auf Peters Routine), aber ich vertraue nicht *ihm*.

Wandeln wir den Fall nun zu einem alternativen Szenario ab: Stellen wir uns vor, Peter sei nicht mein Nachbar, sondern mein Mitbewohner und ich würde ihn am Abend vor dem wichtigen Termin darum bitten, mich am nächsten Tag um 8 Uhr zu wecken. Peter willigt ein dies zu tun. In diesem Fall kann ich zweifellos weiterhin darauf vertrauen, dass mich Peter wecken wird. Darüber hinaus jedoch beschreibt der Fall eine Situation, in der ich auch *Peter vertrauen* kann, dass er mich pünktlich wecken wird.

Diese Beispiele veranschaulichen, dass offensichtlich ein kategorialer Unterschied besteht zwischen dem Vertrauen in eine Person bezüglich einer Handlung auf der einen Seite und dem Vertrauen darauf, dass eine Person diese Handlung ausführt auf der anderen Seite. Einer *Person* P zu vertrauen, sich so-und-so zu verhalten, geht offenbar darüber hinaus, bloß *darauf* zu vertrauen, *dass* diese Person P sich so-und-so verhalten wird. In der Tat begründet diese Differenz eine zentrale Leitunterscheidung innerhalb der philosophischen Diskussion um das Vertrauensphänomen. In der etablierten Terminologie wird sie üblicherweise als Unterschied zwischen dem Vertrauen in eine Person und einem bloßen *Vertrauen-darauf*, oder zuweilen auch einem bloßen *Sich-Verlassen* bezeichnet. Diese Unterscheidung steht auch im Zentrum der hier vorliegenden Arbeit. Sie gilt es nachzuvollziehen und auf den Begriff zu bringen.

In einer ersten vergleichenden Betrachtung der beiden skizzierten Fälle mag einem dabei die Vermutung in den Sinn kommen, dass ich im ersten Falle Peters Verhalten im Prinzip wie einen regelmäßigen und prognostizierbaren Teil meiner Umwelt betrachte. Zu sagen, dass ich darauf vertraue, dass sich Peter wie immer verhalten wird, bedeutet der Form nach zu sagen: Ich vertraue darauf, dass dieses-und-jenes *passieren* wird. Wenn ich im zweiten Fall jedoch *Peter* vertraue, bedeutet dies offenbar etwas kategorial Anderes als dass ich bloß eine prediktive Annahme über Peters zukünftige Verhalten mache. Ich scheine mich also auf Peter nicht einfach nur als einen Teil meiner Umwelt zu beziehen sondern, in einem noch unklaren Sinne, auf Peter *als* einer *Person*. Hier deutet sich an, dass zur akkuraten Erfassung dieses genuin interpersonellen Vertrauens ein grundlegender Wechsel in den Beschreibungskategorien vorgenommen werden muss, und zwar hin zu Kategorien genuiner Interpersonalität.

Trotz der erhöhten Aufmerksamkeit, die das Thema *Vertrauen* in den vergangenen Jahrzehnten erhalten hat – und zwar nicht nur in der philosophischen Diskussion, sondern allgemein in den geistes-, und gesellschaftswissenschaftlichen Disziplinen – scheint der Unterschied zwischen interpersonellem Vertrauen und Vertrauen-darauf noch nicht klar genug

erfasst worden zu sein. Ein gewisser Teil der vertrauenstheoretischen Ansätze, wie etwa die Spieltheorie, sind dabei bereits von ihren theoretischen Grundzügen her blind für diese Unterscheidung. Dort jedoch, wo die Unterscheidung als relevant hervorgehoben wird, treten Probleme dabei auf, diesen klar zu konzeptionalisieren. Zwar lässt sich für das bloße Vertrauen-darauf eine im Wesentlichen unumstrittene Charakterisierung finden; grob gesehen kann man darunter nämlich eine optimistische Haltung angesichts eines Risikos verstehen. Wenn es demgegenüber jedoch darum geht, die Frage zu beantworten, was es heißt, einer Person zu vertrauen, herrscht eine gewisse artikulatorische Unsicherheit vor.

Um die Spezifität interpersonellen Vertrauens festzumachen, werden in der bestehenden Literatur vor allen Dingen zwei Distinktionsmerkmale postuliert. An erster Stelle wird dabei üblicherweise hervorgehoben, dass Vertrauen in eine Person mit einer normativen Erwartung an diese einhergeht. So wurde in vielen philosophischen Beiträgen gezeigt, dass, wenn wir einer Person P vertrauen ϕ zu tun, dies bedeutet, dass wir nicht nur davon ausgehen, dass P ϕ tun wird, sondern auch, dass P ϕ tun *sollte*. Als zweites Charakteristikum interpersonellen Vertrauens weisen Teile der Literatur darauf hin, dass wir im Vertrauen in eine Person, dieser bestimmte Handlungsmotive zuschreiben, aus der heraus sie das Vertrauen erfüllen wird. Einfach gesagt, bedeutet demnach einer Person zu vertrauen immer auch, dass wir dieser *gute, vertrauenswürdige Absichten* beimessen und nicht einfach nur eine ereignisförmige Vorstellung davon haben, was diese Person tun wird.

Beide Ideen zusammen genommen begründen das, was wir als *Standardauffassung interpersonellen Vertrauens* bezeichnen können. Dieser Ansatz, der verschiedene Varianten kennt, ist gut begründet und weist meines Erachtens auch auf notwendige Charakteristika interpersonellen Vertrauens hin. Mit ihm deutet sich auch der angesprochenen Kategorienwechsel an, im Zuge dessen die Vertrauensperson *als Person* in den Blick genommen werden kann. Allerdings werden wir in der Arbeit argumentieren, dass beide genannten Ideen in der hier angedeuteten Form und ohne weitere Qualifizierung nicht *hinreichend* sind ein überzeugendes Distinktionsmerkmal zwischen interpersonellem Vertrauen und einem bloßen Vertrauen-darauf zu bilden. Beide fangen also immer noch *nicht erschöpfend* ein, was es heißt jemandem zu vertrauen.

So bleibt das das meines Erachtens hartnäckigste theoretische Problem in Bezug auf die begriffliche Klärung interpersonellen Vertrauens auch für die Standardauffassung zu bestehen. Dies besteht im Kern darin, dass die Standardauffassung eine rein sozialkalkulatorische Ausdeutung derselben nicht ausschließt. Wir werden dies als *Problem der Sozialkalkulation* bezeichnen. Auf

den Punkt gebracht, ergibt sich dieses aus folgender Überlegung: Wir können uns einen rein sozialkalkulierend eingestellten Akteur vorstellen, der sich auch auf die guten, vertrauenswürdigen Absichten eines Interaktionspartners bezieht, allerdings *allein als Moment in einem instrumentellen Kalkül*. Das gute Motiv der anderen Person fungiert in dieser Perspektive wiederum allein als eine Variable zur Prognose deren Verhaltens. Wenn man darauf entgegnend anführt, dass Vertrauen nach Standardauffassung aber auch mit einer normativen Erwartung an den Vertrauenspartner einhergeht, gibt dieser Hinweis als solcher noch keine klare Antwort darauf, inwiefern damit eine bloß sozialkalkulierende Einstellung ausgeschlossen sein soll.

Es scheint jedoch, dass eine adäquate Theorie interpersonellen Vertrauens eine solche kalkulierende Haltung klar auszuschließen hat, da letztere mit genuinem Vertrauen nicht in Einklang zu bringen scheint. Wie wir noch genauer argumentieren werden, geht interpersonelles Vertrauen in einer solchen kalkulierenden Perspektive nicht auf. Jemandem zu vertrauen stellt eine prinzipiell andere Haltung als eine instrumentalistische, berechnende Einstellung dar. Bei der sozialkalkulierenden Haltung scheint es sich damit allenfalls um eine spezifische, man möchte vielleicht sagen besonders perfide, Form eines Vertrauens-darauf zu handeln, die lediglich die Vertrauenswürdigkeit des Interaktionspartners mit einbezieht.

Das Problem der Sozialkalkulation deckt auf, dass der Wesenskern interpersonellen Vertrauens mit der Standardauffassung noch nicht hinreichend erfasst ist. Demnach ist es nicht hinreichend zu postulieren, dass Vertrauen mit einer normativen Erwartung einhergeht und sich auf ein angenommenes Motiv des Vertrauenspartners bezieht. Damit ist noch nicht erläutert wie genau – nämlich nicht auf sozialkalkulierende Art – sich der Vertrauende auf die andere Person bezieht. Vor diesem Hintergrund wird diese Arbeit für die Hauptthese argumentieren, dass interpersonelles Vertrauen als eine im bestimmten Sinn verstandene Form von *personaler Anerkennung* aufgefasst werden muss.[3]

Anerkennung im hier einschlägigen Sinne ist dabei keine bloß epistemische Kategorie. Anerkennung muss in diesem Zusammenhang vielmehr als ein Respons auf einen *normativen Anspruch der Vertrauenspartnerin* verstanden werden. Was es heißt jemandem zu vertrauen, lässt sich vor dem Hintergrund erläutern, dass von der Vertrauenspartnerin ein normativer Anspruch ausgeht als vertrauenswürdig angesehen und wertgeschätzt zu werden. Vertrauen stellt eine Anerkennung dieses Anspruches und damit eine Anerkennung der anderen, mit einem normativen

3 Dass Vertrauen einen anerkennenden Charakter hat, wurde zuweilen in der bestehenden Literatur schon festgestellt. Im Fokus steht sie auch in Teilen der Literatur zur sogenannten epistemischen Ungerechtigkeit. In der vorliegenden Arbeit soll dies aber als begrifflicher Kern von Vertrauen herausgestellt werden.

Status ausgestatteten, Person dar. In diesem Punkte erweist sich Vertrauen als inkommensurabel mit einer sozialkalkulatorischen Perspektive.

Diese Idee verspricht einen Ansatz bereitszustellen, der die Spezifität interpersonellen Vertrauens erfasst und damit das Problem der Sozialkalkulation dezidiert und nicht bloß *ad-hoc* umgehen kann. Gemäß diesem Ansatz rekurriert interpersonelles Vertrauen zwar auch auf die Vertrauenswürdigkeit der vertrauten Person. Der genaue Zusammenhang wird aber so spezifiziert, dass ein rein sozialkalkulierendes Verständnis dieser Bezugnahme konzeptionell ausgeschlossen ist. Die Vertrauenswürdigkeit der vertrauten Person fungiert schließlich nicht als Moment eines sozialen Kalküls, sondern als eine anzuerkennende Charaktereigenschaft.

Die Tugend dieses anerkennungstheoretischen Ansatzes besteht aber nicht allein darin, den Unterschied zwischen interpersonellem Vertrauen und Vertrauen-darauf auf den Begriff zu bringen. Folgende weitere theoretische Vorzüge derselben werden wir herausstellen:

1) Der anerkennungstheoretische Ansatz erklärt unmittelbar weitere phänomenologische Besonderheiten interpersonellen Vertrauens, wie etwa die Qualität der emotionalen Reaktionen, die wir auf uns entgegengebrachtes Vertrauen oder Misstrauen haben können.

2) Er öffnet den Horizont dafür, die besondere Signifikanz zu erschließen, die Vertrauen im Kontext und zur Gestaltung interpersoneller Beziehungen hat. Hierbei stellt er keine eingeengte Vertrauenskonzeption dar. Entgegen Tendenzen in der Literatur das Vertrauensphänomen eng mit besonderen Beziehungsformen, insbesondere mit persönlichen und affektiven Beziehungsformen wie Freundschaft oder Liebesbeziehungen, zu assoziieren, verspricht der anerkennungstheoretische Ansatz weit genug gefasst zu sein, um der breiteren Vielfalt an denkbaren Vertrauensrelationen – auch unpersönlicher, flüchtiger oder anonymer Natur – gerecht zu werden.

Die Erarbeitung des anerkennungstheoretischen Ansatzes stellt nach der Einleitung den zweiten Teil dieser Arbeit dar. Sie eröffnet eine Perspektive auf interpersonelles Vertrauen als einer bestimmten Art des Verhältnisses zwischen zwei Personen *als Personen*. Es ist diese Betrachtungsweise, die ich die *interpersonelle Seite* des Vertrauensphänomens nennen werde. Dafür wird bewusst die Betrachtung dessen, was ich als epistemische Seite von Vertrauen bezeichne, eingeklammert, die erst im anschließenden dritten Teil der Arbeit adressiert werden soll.

Die epistemische Betrachtungsweise ist insbesondere innerhalb der Diskussion um Sprechervertrauen, der sogenannten Zeugenschaftsdiskussion, vorherrschend, die sich mit dem

Vertrauen in Tatsachenberichte von Sprecherinnen auseinandersetzt. Innerhalb dieser Debatte besteht ein zentrales Desideratum darin zu erklären, wie es möglich sein kann, Wissen über die Welt über die Aussagen anderer zu erhalten. Dies scheint nur möglich, wenn die Hörerin annehmen kann, dass sie es mit einer vertrauenswürdigen Sprecherin zu tun hat. Vor diesem Hintergrund fokussiert die Diskussion vor allem auf die Frage danach, was diese Annahme *epistemisch rechtfertigt*. Welche Gründe kann eine Hörerin dafür haben, davon auszugehen, dass eine Sprecherin tatsächlich vertrauenswürdig *ist*?

Auch unabhängig vom Zeugenschaftskontext ist Vertrauen ein naheliegender Gegenstand dieser epistemischen Fragestellung. Dies liegt im Kern darin begründet, dass Vertrauen intrinsisch immer schon mit einem *epistemischen Anspruch* einhergeht. Einfach gesagt, verstehen wir uns als Vertrauende nicht einfach als arbiträr „irgendjemandem" vertrauend. Wir vertrauen vielmehr, zumindest dem Anspruch nach, nur den Personen, von denen wir annehmen, dass sie tatsächlich vertrauenswürdig *sind.* In diesem Sinne vertrauen wir niemals absolut „blind". Dieser Gedanke wird noch eindrücklicher, wenn man bedenkt, dass wir in Bezug auf Vertrauen und Misstrauen alltäglich *Unterschiede* ziehen: Zu manchen Personen haben wir eine vertrauende Haltung, anderen Personen misstrauen wir. Das Ziehen dieser Unterschiede verstehen wir jedoch nicht als eine bloß willkürliche Entscheidung, sondern als zumindest nicht vollkommen losgelöst von einem vernünftigen Urteil darüber, wer tatsächlich vertrauenswürdig ist und wer nicht. Insofern wir einer Person vertrauen, sprich sie als vertrauenswürdig ansehen, gehen wir also davon aus, dass etwas dafür (oder zumindest nichts ausschlaggebend dagegen) spricht, dass sie auch vertrauenswürdig ist. In dieser Überlegung zeigt sich somit ein innerer, begrifflicher Zusammenhang zwischen Vertrauen und epistemischen Gründen.

Das leitende Forschungsinteresse des dritten Teiles dieser Arbeit besteht nun darin zu verstehen, wie genau dieser Zusammenhang zwischen Vertrauen und seinen epistemischen Gründen zu denken ist. Vertrauen und seine Gründe hängen nämlich auf eine spezifische, vertrauens-typische Weise und nicht einfach *irgendwie* zusammen. Bestimmte mögliche Zusammenhänge zwischen Vertrauen und seinen Gründen sind mit genuinem Vertrauen nämlich gerade unvereinbar.

Diese Diskussion wird zunächst einmal in Bezug auf Vertrauen im Allgemeinen geführt, also nicht speziell in Bezug auf interpersonelles Vertrauen. Dabei ist festzustellen, dass das Verhältnis zwischen Vertrauen und seinen Gründen in *verschiedenerlei* Hinsicht erläuterungsbedürftig ist. Ich schlage deshalb vor, zwischen mindestens zwei solcher Hinsichten zu

unterscheiden[4], die auseinandergehalten werden müssen, in der Diskussion jedoch tendenziell verwischt werden.

1) Zunächst einmal ist Vertrauen in einem basalen begrifflichen Sinne mit dem aktualen und ernsthaften Erwägen von Gründen und Gegengründen für dieses Vertrauen nicht vereinbar. Damit ist gemeint: Man kann nicht *gleichzeitig* vertrauen und ernsthaft sein Vertrauen überdenken. *Insofern* man nämlich vertraut, etwa dass x geschehen wird, *geht man davon aus*, dass x geschehen wird. Auf etwas zu vertrauen bedeutet also diesbezüglich *festgelegt* zu sein. In diesem Sinne ist Vertrauen begrifflich mit Zweifeln und Deliberation inkompatibel.

2) Nichtsdestotrotz *kann* unserem Vertrauen ein Prozess expliziten Überlegens *vorangehen*. Zuweilen überlegen wir erst, ob wir vertrauen sollen und entschließen uns erst daraufhin dazu. In vielen Fällen jedoch vertrauen wir in einem nicht ausdrücklich reflektierten Modus, sondern in präreflexiver und sich praktisch vollziehender Weise. Ich verstehe dies als eine habitualisiert vorliegende Form von Vertrauen, die in mannigfachen alltäglichen Praktiken oder als individueller Habitus des Vertrauens vorliegt. Diese konstituieren einen *selbstverständlichen* Bezug zur Welt, zu einem Selbst und (im Falle interpersonellen Vertrauens) zu Anderen. In dieser Form jedoch kann Vertrauen gerade seinen basalen Selbstverständlichkeitscharakter verlieren, sofern dieses ausdrücklich nach seiner Rechtfertigung befragt wird oder dieses überhaupt thematisch gemacht wird. Mit anderen Worten: das explizite Erwägen von Gründe für Vertrauen kann selbiges in seiner habitualisierten Form als Praxis unterminieren.

Die Darstellung dieser Punkte ist zugleich vorbereitend, um klarer das letztendliche Problem abzugrenzen, das im vierten und letzten Teil dieser Arbeit behandelt werden soll. Dieses Problem besteht darin, wie die interpersonelle und die epistemische Sichtweise auf Vertrauen sinnvoll integriert werden können. Allgemein kann man sagen, dass der Diskurs über interpersonelles Vertrauen tendenziell dichotom geführt wird: Einerseits allein in Hinblick auf seine epistemische Rechtfertigung, andererseits allein in interpersoneller Hinsicht, das heißt mit dem Ziel, die Art der Beziehung zwischen den beiden Vertrauenspartnern phänomenangemessen

4 Man könnte auch noch eine weitere Hinsicht einfügen, die ich in der Diskussion nur *en passant* behandel. Diese bezieht sich auf die Stärke der Gründe, die für Vertrauen erforderlich sind: Insofern wir vertrauen, nehmen wir an, dass eine gewisse Unsicherheit besteht, wir also keine garantierenden Gründe haben. An diesem Punkt unterscheidet sich Vertrauen von anderen epistemischen Einstellungen, wie etwa von Wissen (vgl. Simmel 1992, 393).

auf den Begriff zu bringen.

In der philosophischen Vertrauensdebatte wurde schon mehrfach die Schwierigkeit festgestellt, die epistemische Rechtfertigung von Vertrauen mit seinem genuin interpersonellen Charakter zu versöhnen. Für uns stellt sich dieses Problem in einer spezifischen Variante, nämlich unter dem Vorzeichen der im zweiten Teil erarbeiteten Anerkennungstheorie des Vertrauens dar: Wie kann Vertrauen *als personale Anerkennung* mit seinem Anspruch auf *epistemische Rechtfertigung* und seinem Zusammenhang zu epistemischen Gründen angemessen zusammen gedacht werden? In der Literatur lässt sich diesbezüglich eine in verschiedenen Spielarten ausbuchstabierte Strategie identifizieren. Ich nenne diese den *additiven Lösungsansatz*. Dieser gestaltet sich grundlegend so, dass er zwei methodische Schritte auseinanderhält: Demnach gelte es zum einen eine korrekte Beschreibung der interpersonellen Beziehung der beiden Vertrauenspartner zu liefern (in unserem Fall als einer Anerkennungsbeziehung). Davon *unabhängig* gelte es aber schlichtweg darauf hinzuweisen, dass die Vertrauende auch epistemische Gründe dafür haben muss, um den Vertrauenspartner als vertrauenswürdig anzusehen. Wie wir argumentieren werden scheint diese methodische Trennung aber nicht aufzugehen: Schließlich scheint es Fälle zu geben, in denen die besondere *Art* an epistemischen Gründen, und wie man sich auf diese bezieht, gerade dem Vertrauen in eine Person zuwiderlaufen. Das Problem des additiven Ansatzes besteht im Kern darin, dass er den begrifflichen Bezug zwischen der anerkennenden Seite von Vertrauen und epistemischen Gründen nicht herausarbeitet. Damit lässt er unspezifiziert, auf welche Arten von epistemischen Gründen man sich als Vertrauender beziehen kann und erlaubt damit auch solche Fälle der epistemischen Absicherung von Vertrauen, die jenes gerade konterkarieren.

Angesichts dieses Defizits des additiven Ansatzes werden wir davon ausgehen, dass sich die interpersonelle und die epistemische Seite von Vertrauen nicht vollständig voneinander trennen und einfach aufaddieren lassen. Vielmehr muss ein *intrinsischer* analystischer Zusammenhang zwischen der Beschreibung der interpersonellen Seite von Vertrauen mit seiner epistemischen Seite bestehen. Diesen Zusammenhang werden wir ausgehend von einer weitergehenden Analyse des Anerkennungsbegriffes aus dem zweiten Teil aufzeigen. Das Ergebnis dieser Analyse wird sein, dass der Anerkennungsbegriff von sich aus bereits auf epistemische Zusammenhänge verweist. Ansonsten könnte Anerkennung gewissermaßen kein *Ernsthaftigkeitscharakter* zukommen. Damit werden wir aufzeigen, dass sich die interpersonelle Seite von Vertrauen notwendigerweise auch auf die epistemische Seite erstreckt. Erst ausgehend von einem klaren Verständnis der Relation der

beiden Vertrauenspartner lässt sich somit der genaue Stellenwert epistemischer Gründe im Kontext interpersonellen Vertrauens angemessen verorten.

1.2 Methodische Bemerkungen: Orientierung am gewöhnlichen Begriffsverständnis

Grundlegend gesprochen steht im Zentrum dieser Arbeit die Frage nach der begrifflichen Klärung eines Terminus. Es geht im Grunde um die Frage „Was ist interpersonelles Vertrauen?“ oder „Was bedeutet es jemandem zu vertrauen?“. Diese lässt sich in die Reihe der für die Philosophie grundlegenden „Was-ist-überhaupt-X-Fragen“ einreihen, wie etwa „Was ist überhaupt Wissen/eine Tugend/Legitimität etc?“. Entscheidend ist zu sehen, dass, insofern es sich dabei um begriffliche Frage handelt, diese grundsätzlich nicht mithilfe neuer empirischer Informationen beantwortet werden kann. Vielmehr zielt sie darauf ab, ein bereits vorausgesetztes Begriffsverständnis zu *explizieren*.

In einem vortheoretischen Sinne wissen wir bereits, was Vertrauen in eine Person bedeutet. Wir haben diesbezüglich ein ursprüngliches Phänonemverständnis. Dieses nun spiegelt sich auch in unserem Verständnis und der Verwendungsweise der alltagssprachlichen Formulierung „jemandem vertrauen“ wider. Diese alltagssprachlich verankerte Redeweise zu verstehen, bedeutet deshalb auch das Phänomen zu verstehen, auf das es referiert. Wie insbesondere Autor:innen, die der sogenannten Philosophie der gewöhnlichen Sprache zugerechnet werden, verdeutlicht haben, fällt deshalb die Klärung eines im lebensweltlichen Alltag situierten Begriffes mit der Klärung des damit referierten Phänomens zusammen. In diesem Sinne ist die Klärung von Vertrauen *de dicto* äquivalent mit der Klärung von Vertrauen *de re*. Deswegen ist der Philosophie der gewöhnlichen Sprache, entgegen eines verbreiteten Missverständnisses, auch keine realitätsvergessene Sprachfokussierung vorzuwerfen. Gegen dieses Vorurteil gerichtet, stellte Stanley Cavell entsprechend prägnant fest: „the philosophy of ordinary language is not about language, anyway not in any sense in which it is not also about the world. Ordinary language philosophy is about whatever ordinary language is about.“[5]

5 Cavell 2015, 89. Beeindruckend klar argumentiert auch Oswald Hanfling gegen den Vorwurf der vermeintlichen bloßen Sprachfokussierung in der Philosophie der gewöhnlichen Sprache: „Nevertheless the *questions* ‘What is X?’ and ‘What does “X” mean?’ may be one and the same. Let us consider how an enquiry about ‘the thing itself’ is to be conducted. Asked to say what knowledge or courage are, how am I to proceed? How should I investigate ‘the thing itself’, say knowledge, as distinct from the word? It is not as if knowledge were some kind of stuff, so that we might ask what it consists of, as distinct from asking what the word means“ (Hanfling 2000, 17).

Dieser grundlegenden Einsicht in methodischen Orientierungsfragen der Philosophie folgend, beziehen wir uns in der philosophischen Untersuchung interpersonellen Vertrauens wesentlich auf unser alltagssprachlich fundiertes Begriffsverständnis.[6] Unser gewöhnliches Begriffsverständnis interpersonellen Vertrauens markiert gewissermaßen den Orientierung gebenden, methodischen Fixpunkt unseres Klärungsversuches. Die eigentliche philosophische Herausforderung besteht nun jedoch darin, in rekonstruktiver Arbeit adäquat *darzulegen*, wie genau wir den Begriff interpersonellen Vertrauens immer schon verstehen. In diesem Sinne ist diese Arbeit zwar methodisch an unserem gewöhnlichen Vertrauensbegriff orientiert, aber nicht in einem empirisch-linguistischen Sinne. Schließlich versteht sie ihre eigentliche Anstrengung wesentlich in einer *Explizierungs*arbeit und nicht in dem *Entdecken* von etwas noch nicht Gewusstem. Indem sie dem Alltagsverständnis nachspürt, soll sie freilegen, was die begriffliche Struktur interpersonellen Vertrauens ausmacht.

Auf einer spezifischeren methodischen Ebene besteht dabei eine übliche Vorgehensweise im Betrachten von Beispielfällen, an denen die Struktur unseres Alltagsbegriffes interpersonellen Vertrauens aufgezeigt werden kann. Um es aber noch einmal zu betonen, sind diese Beispiele nicht zu dem Zweck gedacht neue *empirische* Informationen über das Phänomen interpersonellen Vertrauens bereitzustellen. Die Beispielfälle stellen vielmehr den Prüfstein dar, an dem sich die philosophischen Explikationsbemühungen messen lassen müssen. Sollte eine philosophische Theorie des Vertrauens zulassen, dass bestimmte Fälle als Exemplifizierung von Vertrauen ausgewiesen werden müssten, obwohl dies unserem vortheoretischem Begriffsverständnis radikal widerspricht, stellt sich dies als Problem für diese Theorie dar.

Das Aufdecken einer begrifflichen Struktur soll hier im Übrigen nicht mit dem Angeben einer Definition in einem bestimmten engen Sinne verwechselt werden. Uns geht es nicht bloß

6 Grundlegend orientiert sich diese Arbeit also an dieser groben methodischen Ausrichtung der Philosophie der gewöhnlichen Sprache (*Ordinary language philosophy*), wie man sie etwa in den Schriften des späten Wittgensteins exemplifiziert findet. Unter dem Titel der Philosophie der gewöhnlichen Sprache soll hier allerdings allein eine *grobe methodische Orientierung* philosophischer Untersuchungen verstanden werden, die versucht, bloße Spekulationen über einen Begriff oder das Wesen einer Sache zu vermeiden. Unser „alltagssprachliches Verständnis“ markiert demnach gewissermaßen den methodischen Ankerpunkt einer philosophischen Untersuchung. Am prägnantesten finde ich diese Idee ausgedrückt im §106 der Philosophischen Untersuchungen von Wittgenstein: „Wenn die Philosophen ein Wort gebrauchen - »Wissen«, »Sein«, »Gegenstand«, »Ich«, »Satz«, »Name« - und das *Wesen* des Dings zu erfassen trachten, muß man sich immer fragen: Wird denn dieses Wort in der Sprache, in der es seine Heimat hat, je tatsächlich so gebraucht? -*Wir* führen die Wörter von ihrer metaphysischen, wieder auf ihre alltägliche Verwendung zurück“ (Wittgenstein 2003a, 82, § 116). Davon unabhängig wollen wir uns damit nicht auf ein bestimmtes *Schema* festlegen, wie *genau* ein Begriff der Alltagssprache zu analysieren ist, etwa nur in Form der Angabe seiner notwendigen und hinreichenden Anwendungsbedingungen (kritisch dagegen jedoch: Wittgenstein 2003a, § 41-43, § 47). Eine ausführliche Darstellung und auch Diskussion der üblichen, jedoch meist auf Missverständnissen beruhenden Einwände gegenüber der Grundidee der Philosophie der gewöhnlichen Sprache, bietet: Hanfling 2000.

darum definierende Elemente, die *definientia,* interpersonellen Vertrauens aufzuzählen – etwa in der Form der Angabe seiner notwendigen und hinreichenden Anwendungsbedingungen eines Begriffes, die ein Standardschema der traditionellen analytischen Philosophie zur Begriffsanalyse darstellte. Dass wir diesen Weg nicht einschlagen werden, begründet sich dabei nicht nur aus einer philosophiehistorisch motivierten Skepsis darüber, wie weitreichend sich die Bedeutung von Begriffen überhaupt erschöpfend durch die Angabe der Summe seiner definierenden Elemente analysieren lässt. Um ein klassisches Beispiel anzuführen, sei hier etwa an den Versuch erinnert, die Bedeutung eines Begriffes wie „gut“ zu verstehen. Wie Peter Geach gezeigt hat, lässt sich dieser Begriff nicht als Bündel etwaiger objektstufiger Eigenschaften analysieren. Vielmehr ist er nur vor dem Hintergrund eines bestimmten normativen Formverständnisses nachvollziehen, das es genauer zu erläutern gilt: Das Attribut „gut“ signifiziert demnach, dass etwas die evaluativen Standards seines eigenen Begriffes, unter den es fällt, über einem bestimmten Maß erfüllt.[7] Zur Erläuterung der Bedeutung des Attributes „gut“ muss dieser strukturbezogene Zusammenhang erläutert werden, statt vermeintliche definierende Elemente desselben aufzuzählen.

Über diese begründete Skepsis hinaus geht es uns in dieser Arbeit aber auch prinzipiell nicht um die Erarbeitung einer bloßen Definition von Vertrauen, sondern darum, die weiterreichende und begrifflich komplexere Struktur zu erläutern, die der Vertrauensbegriff impliziert.[8] So ist die leitende methodische Annahme dieser Arbeit, dass man nur insofern versteht was interpersonelles Vertrauen bedeutet, als man auch die Beziehung zwischen Vertrauenspartnern und den Zusammenhang zwischen Vertrauen und seinen epistemischen Gründen versteht. Eine begriffliche Analyse von Vertrauen muss demnach auch eine Erörterung dieser interpersonellen und epistemischen Verflechtungen des Vertrauensbegriffes vornehmen, die eben über eine bloße Definition im engeren Sinne hinausgeht. Was wir damit am Ende erarbeiten, könnte man alternativ auch eine umfassende *Konzeption* von Vertrauen nennen.

In ebendiesem Sinne fasse ich auch den im Titel dieser Arbeit platzierten Wittgenstein‘schen Begriff der *Grammatik* auf. Unter dem Begriff der Grammatik fasste Wittgenstein am einfachsten ausgedrückt die Menge aller Regeln, die den Gebrauch von Sprache

7 Geach 1956. Um eine prägnante Erläuterung dieses Gedankens anzuführen: „if I do not know what hygrometers are for, I do not really know what 'hygrometer' means, and *therefore* do not really know what 'good hygrometer' means“ (ebd.: 38). Zur Abgrenzung dieser Art der Begriffsanalyse von der Analyse nach der extensionalen Logik, siehe auch: Hoffmann 2015.

8 Diverse verstreute Wittgenstein‘sche Bemerkungen sind auch ausdrücklich kritisch gegen dieses Vorgehen gerichtet, etwa in: Wittgenstein 1969. Man beachte etwa folgende Bemerkung: „The idea that in order to get clear about the meaning of a general term one had to find the common element in all its applications has shackled philosophical investigation“ (ebd.: 19–20).

regulieren.[9] Dies umfasste allerdings nicht nur die auf Konventionen beruhenden syntaktischen Regeln, die im üblichen linguistischen Sinne als Grammatik verstanden werden. Grammatische Regeln in Wittgensteins weitem Sinne geben schlichtweg den *sinnvollen Gebrauch* eines Wortes überhaupt an und schließen unsinnige Aussagen aus.[10] Deshalb schlug Wittgenstein der Grammatik des Gebrauchs der Farbwörter etwa auch eine Regel wie „Verschiedene Farben können nicht zugleich am selben Ort sein“ zu. Zur Grammatik des Wortes Mensch gehörte eine Aussage wie „Menschen können zum Mond fliegen.“[11] Wer nämlich sagt, ein Tisch sei auf seiner gesamten Oberfläche zugleich grün und rot, sage etwas Unsinniges. Ergo verstoße er gegen die Grammatik von Farbwörtern. In diesem weiten Sinne soll es auch hier um die Grammatik von Vertrauen gehen: Es soll schlichtweg wiedergegeben werden, wie wir adäquat und sinnvoll über Vertrauen und seine Zusammenhänge sprechen. Dies beinhaltet auch die Aufgabe, darzulegen, wie wir sinnvoll über den Zusammenhang von Vertrauen und seinen Gründen sprechen, oder wie wir über die Beziehung zweier Personen sprechen, die sich vertrauen.[12]

Auch wenn es also im Wesentlichen darum geht, unser immer schon bestehendes Vertrauensverständnis zu rekonstruieren, ist diese rekonstruktive Tätigkeit keinesfalls trivial. Das sieht man allein an den bereits vorliegenden, jedoch problembehafteten Versuchen, eine Konzeption interpersonellen Vertrauens vorzulegen. Problembehaftet sind diese insofern sie unser gewöhnliches Verständnis interpersonellen Vertrauens nur ausschnittsweise oder verzerrt wiedergeben. Das Problematische an den kritisierten Positionen beruht dabei stets auf einseitigen Abstraktionen: So vermögen sie es nur einem Teilaspekt unseres Vertrauensverständnisses Rechnung zu tragen, aber nicht ein vollständiges Bild desselben zu zeichnen. Demgegenüber ist ein Anspruch dieser Arbeit interpersonelles Vertrauen nicht einseitig zu modellieren, sondern vielmehr die komplexen begrifflichen Zusammenhänge zwischen seinen verschiedenen Aspekten darzulegen. Dabei werden auch die komplexen begrifflichen Zusammenhänge zwischen der normativen, emotionalen und kognitiven Dimension von Vertrauen entfaltet werden.

Wenn im Laufe der Arbeit eine bestimmte Vertrauenskonzeption kritisiert wird, ist damit

9 Siehe etwa: Wittgenstein 2003b, 302-305.
10 Vgl. Moyal-Sharrock 2020.
11 Vgl. ebd., 48.
12 Mit Wittgensteins weiten Grammatik-Begriff wird erneut deutlich, dass der Vorwurf der Realitätsvergessenheit der Philosophie der gewöhnlichen Sprache nicht zutrifft. In der alltagssprachlichen Grammatik von Wörtern zeigen sich schließlich immer auch Fakten über die Wirklichkeit indirekt verdichtet (vgl. Moyal-Sharrock 2020). Bestimmte Aussagen (über Farben etwa) sind sinnvoll, da sie den wirklichen Verhältnissen entsprechen. Die Frage nach der Grammatik eines Begriffes und dem Wesen eines Phänomens ist deshalb nicht strikt voneinander zu trennen. In diesem Sinne verstehe ich auch die Aussage Wittgensteins: „Das Wesen ist in der Grammatik ausgesprochen“ (Wittgenstein 2003a, § 189, § 371).

natürlich nicht automatisch gemeint, dass diese unter allen Umständen zu *verwerfen* ist. Innerhalb von Teildisziplinen mag eine solche vielmehr immer noch ihren jeweiligen Zweck erfüllen – und zwar womöglich sogar *gerade* da es sich um abstrahierende Konzeptionen handelt, die zur Beantwortung bestimmter wissenschaftlicher Fragestellungen besonders fruchtbar und heuristisch handhabbar sind. Beispielsweise mag die Vertrauenskonzeption der Spieltheorie ein probater Ansatz sein, um einen Teilaspekt des Verhaltens von Akteuren innerhalb spezifischer Handlungskontexte nachzuvollziehen und ausreichend akkurat prognostizieren zu können. Aus der philosophisch-klärenden Perspektive weisen wir jedoch nach, inwiefern das in seiner Modellierung vorliegende Verständnis von Vertrauen von einem komplexeren alltagssprachlichen Verständnis abweicht und damit also nicht gleichzusetzen ist.

Neben den in den Einzelwissenschaften verwendeten Konzeptionen von Vertrauen beziehen wir uns jedoch überwiegend auf die bereits in der philosophischen Literatur vorliegenden Klärungsversuche unseres gewöhnlichen Verständnisses. Auch diese werden wir in der Regel als einseitige Abstraktionen kritisieren. Sie abstrahieren von dem, was ich als den eigentlichen Kern interpersonellen Vertrauens ansehe, nämlich dessen Charakter als Anerkennung.

Zum Teil scheint mir eine Quelle des einseitigen Charakters der vorliegenden Rekonstruktionen darin zu liegen, dass zum Teil ein gewisses Vorverständnis der Art der Beziehung zwischen den Vertrauenspartnern vorausgesetzt und als eigentliche oder zumindest prototypische Vertrauensbeziehung ausgewiesen wird. Genauer gesagt wird Vertrauen häufig als etwas verstanden, das im Rahmen von intimen Beziehungen, wie Freundschaften, Liebesbeziehungen oder der Beziehung zwischen Eltern und Kindern, zu seiner eigentlichen Geltung komme.[13] Darüber hinaus sollte allerdings nicht aus dem Blick geraten, dass Vertrauen ebenso in formellen oder gar anonymen Beziehungen auftreten kann, etwa wenn wir Personen in ihrer gesellschaftlichen Funktion vertrauen. In beiden Extremen dieses Spektrums aber sprechen wir von Vertrauen, auch wenn sich dieses natürlich unterschiedlich ausgestaltet. Diese verschiedenen Konfigurationen, die Vertrauen je nach Beziehungskontext haben kann, werden wir zwar auch in dieser Arbeit beleuchten. In erster Linie geht es uns jedoch darum einen möglichst *allgemeinen* Begriff interpersonellen Vertrauens auszuarbeiten, der verschiedene Ausprägungsformen desselben zu überbrücken vermag. Damit grenzt sich unser primäres Vorhaben auch von den Versuchen ab, eine Taxonomie *unterschiedlicher* Vertrauenstypen aufzustellen, die sich nicht nur in den empirischen Einzelwissenschaften, sondern auch im philosophischen

13 Vgl. Hartmann (2011, 27) zu dieser Diagnose. Die Konzeption von Bernd Lahno etwa scheint sich prototypisch am Fall der Freundschaft zu orientieren (siehe insbesondere: Lahno 2002, 132). Ebenso: McLeod 2002, 15.

Vertrauensdiskurs finden lassen.[14] Unser Vorhaben besteht schließlich in erster Linie darin, allgemein zu klären, was es *überhaupt* heißt jemandem zu vertrauen. Die in dieser Arbeit entwickelte Konzeption von Vertrauen als Anerkennung hat entsprechend nicht den Anspruch, einen (gleichgültig wie bedeutsamen) Sonderfall einer Vertrauensbeziehung zu behandeln, sondern vielmehr den allgemeinen begrifflichen Kern interpersonellen Vertrauens überhaupt freizulegen.

Zuletzt sei noch klargestellt, dass wir uns zwar, wie gesagt, am gewöhnlichen Verständnis interpersonellen Vertrauens orientieren werden. Dies soll allerdings nicht bedeuten, dass die *Wörter* unserer Alltagssprache als solche bereits für alle Zwecke immer schon optimal seien. Im Zuge ihrer Untersuchung kann sich vielmehr herausstellen, dass diese *Mehrdeutigkeiten* aufweisen, was einen ihrer potenziellen Mängel darstellt.[15] Dies stellt meines Erachtens auch Stolpersteine bei der Klärung des Vertrauensbegriffes bereit. Auf einen davon werden wir etwa direkt im nächsten Kapitel genauer eingehen: So wird die Vokabel „Vertrauen“ zuweilen alltagssprachlich als Bezeichnung einer bestimmten *Handlung* verstanden, genauer gesagt als bloßes Eingehen eines Risikos. Andererseits bezieht sich das Wort zuweilen aber auch auf eine spezifische *Haltung* eines Akteurs. Aufgrund dieser Ambiguität können sich unterschiedliche Konzeptionen von „Vertrauen“ ergeben – oder besser gesagt: Unterschiedliche Konzeptionen dessen, was nominal identisch mit der Vokabel „Vertrauen“ belegt wird. Ein simpler methodischer Kniff um damit umzugehen, kann darin bestehen, schlichtweg auf die Ambiguität des Signifikanten „Vertrauen“ hinzuweisen und zwischen seinen zwei verschiedenen Bedeutungen zu unterscheiden. Aus solchen Gründen werde ich anlassbezogen immer wieder Unterscheidungen zwischen verschiedenen technischen Vertrauensbegriffen einführen, und zwar immer dort, wo dies sachlich geboten ist. Insgesamt wird damit im Laufe der Arbeit ebenfalls eine Art Taxonomie unterschiedlicher Vertrauens*begriffe* erarbeitet, die sich hinter dem alltagssprachlichen *Wort* „Vertrauen“ verbergen. Dies soll helfen potenzielle Missverständnisse auf sprachlicher Ebene zu vermeiden.

Zusammenfassend können wir das methodische Selbstverständnis dieser Arbeit in folgenden Thesen festhalten:

14 Als Beispiele für Ansätze diesbezüglich in der philosophischen Literatur, siehe insbesondere: Baier 1986; Baker 1987, 1-3; Lahno 2002, insbes. Kap.6; Smolkin 2008, 441-445. Kritisch gegen das damit oft verbundene Vorhaben, philosophische Probleme mit einer Art „Taschenspielertrick“ aufzulösen: Kaminski 2017. In ähnlicher Weise kritisch gegenüber der Taxonomisierungstendenz im Vertrauensdiskurs: Miller 2003.

15 Für eine Auflistung und Erläuterung weiterer Mankos der Alltagssprache, siehe: Hahn 2018, 36-42. Auch dieser Punkt widerspricht meiner Auffassung nach gar nicht dem Ansatz der Philosophie der gewöhnlichen Sprache. Das wäre ebenfalls ein Missverständnis, denn auch wenn man zur Behebung der Ambiguität technische Termini einführt, müssen diese wiederum mithilfe einer bereits vorhandenen und allgemein geteilten gewöhnlichen Sprache erläutert werden. Ebenso können sie Teil der gewöhnlichen Sprache werden, wenn sie eine Selbstverständlichkeit werden. Vgl. Hanflings Bemerkung: „the remedy would still be available within ordinary language: one could ask the speaker what he means and he should be able to tell us.“ (2000, 2).

1. Im Kern geht es in der Arbeit um die Frage, was es heißt jemandem zu vertrauen. Auf der grundlegendsten Ebene orientiert sich diese an unserem bereits bestehenden gewöhnlichen Begriffsverständnis. Dieses gilt es (unter Zuhilfenahme von Anschauungsbeispielen) zu explizieren und in eine kohärente Vertrauenskonzeption zu bringen.
2. Bereits vorhandene Konzeptionen sollen als einseitige Abstraktionen kritisiert werden. Ihnen gelingt es nicht vollumfänglich unser gewöhnliches Begriffsverständnis wiederzugeben. Dass diese in dieser Form zu kritisieren sind, soll nicht bedeuteten, dass sie nicht auch innerhalb von Einzelwissenschaften eine relevante wissenschaftliche Funktion haben können.
3. Die genaue Analyse des Begriffs interpersonellen Vertrauens folgt keinem dogmatisch vorgegebenen Analyseschema, insbesondere nicht in der Form der Angabe von notwendigen und hinreichenden Bedingungen. Die Angabe einer Vertrauens*konzeption* ist vielmehr komplexer: Sie verlangt eine Erläuterung weitreichender begrifflicher Zusammenhänge, etwa der Rolle von Vertrauen innerhalb einer Beziehung zwischen zwei Personen oder dem genauen Zusammenhang von Vertrauen mit seinen Gründen. Diesen weiten Sinn der begrifflichen Erläuterung belege ich mit dem Wittgenstein‘schen Begriff der *Grammatik*.
4. Die grundlegende Orientierung an der gewöhnlichen Sprache verbietet nicht, die *Wörter* derselben anlassbezogen als unklar (etwa als mehrdeutig) zu kritisieren und durch präzisere technische Termini zu ersetzen.

1.3 Das Vorgehen im Einzelnen

Im Folgenden zweiten Kapitel dieses ersten vorbereitenden Teils der Arbeit werden wir eine erste, maximal allgemein gehaltene Analyse des Vertrauensbegriffes vornehmen. Diese Analyse bezieht sich auf den generischen Vertrauensbegriff, von dem interpersonelles Vertrauen nur eine Subkategorie darstellt. Vertrauen wird dabei in einem ersten grundlegenden Schritt als eine spezifische Haltung ausgewiesen. Diese mag sich zwar auch in bestimmten Vertrauensakten manifestieren, ist aber nicht mit diesen gleichzusetzen. Anschließend werden zwei notwendige Elemente von Vertrauen herausgearbeitet: Erstens ein gewisses Risikobewusstsein und zweitens

ein Moment von Zuversicht angesichts dieses Risikos.

Im Weiteren soll interpersonelles Vertrauen oder Vertrauen in eine Person als eine spezifische Unterkategorie des weiten oder generischen Vertrauensbegriffes in den Blick genommen werden. Es handelt sich dabei um eine Form von Vertrauen *sui generis*. Die Spezifität derselben herauszuarbeiten ist Ziel des folgenden zweiten Teils der Arbeit. Den Einstieg dazu bildet eine kritische Auseinandersetzung mit der Rational Choice- bzw. Spieltheorie des Vertrauens. Wir werden nachweisen, dass diese höchstens mit einem Begriff generischen Vertrauens zu operieren vermag und das eigentliche Charakteristikum interpersonellen Vertrauens in der Betrachtungsperspektive der Rational Choice-Theorie keinen Platz findet. Im Zuge dessen heben wir in einer ersten Annäherung zwei häufig genannte Besonderheiten interpersonellen Vertrauens hervor, für die die Rational Choice-Theorie kategorial blind ist. Dies ist zum einen die normative Erwartung des Vertrauenden an den Vertrauenspartner – ein Aspekt, der sich schon aus einem frühen und für die Vertrauensdiskussion wegweisenden Aufsatz von Annette Baier herausarbeiten lässt. Zum anderen wird sich zeigen, dass das Vertrauen in eine Person dieser eine vertrauenswürdige Einstellung unterstellt. Beide Punkte legen die Sichtweise nahe, dass sich interpersonelles Vertrauen nicht nur durch seinen Gegenstandsbereich auszeichnet – eben als eine Vertrauenskategorie, die sich auf andere Personen richte – sondern, dass sich damit auch etwas an der Form ändert in der sich der Vertrauensgeber auf den Vertrauensnehmer bezieht.

Das vierte Kapitel zielt darauf ab, die beiden genannten Distinktionsmerkmale interpersonellen Vertrauens besser zu verstehen, indem es sich der Axiologie von Vertrauen widmet. So verspricht ein klareres Verständnis des Wertebezuges von Vertrauen einen theoretischen Ausgangspunkt dafür zu liefern, nachzuvollziehen, worauf sich, erstens, die normative Erwartung des Vertrauenden bezieht und wie genau sich, zweitens, die, dem Vertrauensnehmer unterstellte, vertrauenswürdige Einstellung charakterisieren lässt. Die tatsächlichen Zusammenhänge sind hierbei allerdings nicht so trivial. Wie wir zeigen werden besteht in der axiologischen Betrachtung von Vertrauen eine mögliche Gefahr darin, die normative Beziehung zwischen Vertrauensgeber und -nehmer fehl zu charakterisieren.

Wie wir argumentieren werden, muss letztere aber ins Zentrum gerückt werden, wenn die Natur der normativen Erwartung und die vertrauenswürdige Haltung des Vertrauenspartners angemessen verstanden werden sollen. Dies soll im fünften Kapitel geschehen. Hier werden wir die normative Relation der beiden Vertrauenspartner mit Hilfe einer Erweiterung des Konzeptes der sogenannten bipolaren Obligationen von Michael Thompson beschreiben. Wir werden darauf

aufbauend aufzeigen, dass Vertrauen notwendigerweise stets in einem normativen Kontext situiert ist. Vor diesem Hintergrund soll auch dafür argumentiert werden, dass Vertrauenswürdigkeit in einem bestimmten Sinne als eine praktische Tugend aufzufassen ist.

Das sechste Kapitel stellt das Kernstück des zweiten Teils der Arbeit dar. In diesem werden wir aufzeigen, dass es zur vollständigen Beschreibung interpersonellen Vertrauens weder ausreicht, bloß die normative Situation zwischen den beiden Vertrauenspartnern zu beschreiben, noch, dass es genügt hinzuzufügen, dass sich das Vertrauen auf die Vertrauenswürdigkeit des Vertrauenspartners bezieht. Beide Momente werden üblicherweise als Charakteristika interpersonellen Vertrauens herangezogen, weshalb wir hier auch von der sogenannten Standardauffassung interpersonellen Vertrauens sprechen werden. Zumindest ohne weitergehende Qualifizierung stellt sich für diese jedoch das Problem, das wir als Problem der Sozialkalkulation bezeichnen werden. Demnach lässt die Standardauffassung Fälle bloß sozialkalkulierender Einstellung zu, die einer genuin vertrauenden Einstellung jedoch geradezu entgegenstehen. In kritischer Auseinandersetzung mit der Zeugenschaftstheorie Richard Morans werden wir daraufhin die Idee motivieren, dass Vertrauen in eine Person wesentlich in einer anerkennenden Haltung besteht. Anschließend sollen die weiteren Vorzüge dieser Vertrauenskonzeption herausgestellt werden. Dazu gehört, dass es einen theoretischen Zugriff auf das Vertrauensphänomen zulässt, der dessen Stellenwert im Kontext von und zur Gestaltung interpersoneller Beziehungen erläutert.

Der darauffolgende dritte Teil dieser Arbeit wird sich mit der epistemologischen Struktur von Vertrauen genauer auseinandersetzen. Im siebten Kapitel sollen dafür einige grundlegende begriffliche Zusammenhänge geklärt werden. Dies bezieht sich zunächst wieder auf Vertrauen im generischen Sinne. Dabei werden wir zum einen für die sogenannte kognitivistische Position in der Vertrauenstheorie argumentieren, gemäß derer Vertrauen mit einer bestimmten Überzeugung einhergeht. Anschließend wird dargelegt, dass Vertrauen aufgrund seiner kognitiven Dimension begrifflich mit epistemischen Gründen verbunden ist, also mit solchen Gründen, die dafür oder dagegen sprechen, dass ein bestimmter Sachverhalt zutrifft. Demnach ist Vertrauen aus der Sicht eines Vertrauenden dem Anspruch nach hinreichend epistemisch begründet. Allerdings lässt sich im Anschluss darauf direkt anmerken, dass Vertrauen in einem bestimmten trivialen Sinne dennoch nicht mit der *Deliberation* über epistemische Gründe vereinbar ist. Vertrauen stellt nämlich eine festgelegte Haltung dar, das heißt: Insofern man darauf vertraut, dass x passieren wird, kann man *nicht zugleich ernsthaft* daran zweifeln oder reflektieren, ob x passieren wird.

Diese *triviale* Disjunktion von Vertrauen und Zweifeln bzw. Überlegen ist zu unterscheiden

von einer anderen Hinsicht, in der sich Vertrauen und Zweifel ausschließen. Erstere besagt schließlich nur, dass insofern man vertraut, man nicht zweifeln kann. Dies schließt nicht aus, dass man etwa anfänglich zweifeln mag, anschließend die eigene Skepsis jedoch überwinden kann, um zu einer genuin vertrauenden Einstellung zu gelangen. Vielfach wurde jedoch festgestellt, dass es zumindest bestimmte Kontexte gibt, in denen die Tatsache, dass man überhaupt ausdrücklich darüber reflektiert, ob man vertrauen soll oder nicht, bereits Vertrauen untergräbt oder verunmöglicht. Diese Intuition soll im achten Kapitel präziser geklärt werden. Es soll der Frage nachgegangen werden, was uns diese Intuition über die Beschaffenheit von Vertrauen verrät. Die These, die wir entwickeln werden lautet, dass Vertrauen hierbei implizit als habitualisierte Praxis aufgefasst wird. Vor dem Hintergrund des Verständnisses von Vertrauen als einer habitualisierten Praxis oder als individueller Habitus kann plausibilisiert werden, inwiefern Vertrauen in dieser Form ausdrücklichem Überlegen entgegensteht.

Der abschließende vierte Teil der Arbeit nimmt letztendlich den Zusammenhang zwischen der interpersonellen und epistemischen Seite von Vertrauen in den Fokus. Vielfach wurde bereits festgestellt, dass beide Perspektiven auf Vertrauen tendenziell dichotom nebeneinanderstehen. Ein bei verschiedenen Autor:innen, etwa bei Richard Moran, identifizierbarer Versuch beide Seiten zusammen zu bringen, besteht darin sie einfach additiv zusammen zu denken: Schematisch gesprochen, müsse die interpersonelle Beschreibung von Vertrauen lediglich noch mit dem Hinweis ergänzt werden, dass Vertrauen auch eine epistemische Basis braucht. Wie wir argumentieren werden, übersieht dieser Ansatz jedoch den inhärenten Zusammenhang zwischen beiden Seiten. Dieser zeigt sich etwa darin, dass bestimmte epistemische Gründe nicht wirklich mit dem Vertrauen in eine Person vereinbar sind, da sie gerade dem anerkennenden Charakter von Vertrauen widersprechen – ein Problem, das ich das Problem der defizienten oder pathologischen Vertrauensfälle nennen werde. Ein alternativer Vorschlag, die interpersonelle und epistemische Seite von Vertrauen miteinander in Einklang zu bringen stammt von Edward Hinchman und bezieht sich auf die *entitlement*-Theorie in der Zeugenschaftsdebatte. Wie wir in einem Exkurs zeigen werden, schafft es diese aber nur scheinbar die problematischen Fälle defekten Vertrauens zu umgehen. Der anschließend entwickelte Gegenvorschlag nimmt eine genauere Analyse des Anerkennungsbegriffes aus dem zweiten Teil der Arbeit vor, indem es dessen inhärenten epistemischen Zusammenhang herausstellt. Vor dem Hintergrund dieser Analyse kann der Zusammenhang zwischen Vertrauen als Anerkennung und epistemischen Gründen präziser qualifiziert werden. In einem letzten Unterkapitel der Arbeit setzen wir uns abschließend mit

Überlegungen auseinander, die drohen, Vertrauen letztendlich doch als ein epistemisch irrationales Phänomen auszuweisen. Diese Überlegungen setzen bei dem Gedanken an, dass wir zuweilen nicht aufgrund einer nüchternen Betrachtung von Evidenzen erst dazu kommen zu vertrauen oder zu misstrauen, sondern, dass wir umgekehrt Evidenzen stets bereits im Lichte unserer Vertrauens- und Misstrauensbeziehungen bewerten. Dadurch mag es so aussehen als wäre unser Vertrauen nur scheinbar rational, würde sich aber in Wirklichkeit auf eine problematische zirkuläre Weise selber begründen. Diese Schlussfolgerung muss jedoch relativiert werden. Richtig ist vielmehr nur, dass Evidenzen üblicherweise einen rationalen Spielraum ihrer Interpretation zulassen. Wir mögen deshalb in unseren Einschätzungen, wem man vertrauen darf und wem nicht, voneinander abweichen, ohne dass dies zwangsläufig hieße, dass einer von uns damit irrational handle. Mehr noch werde ich als letzten Gedanken andeuten, dass ein solcher Spielraum gar notwendig ist, um verständlich zu machen, wie Vertrauen eine Art Sich zu Verhalten darstellt, mit der die Beziehung zur anderen Person aktiv gestaltet wird.

2. Der generische Begriff des Vertrauens: Eine erste Annäherung

2.1 Eine formale Definition: Vertrauensakt und vertrauende Haltung

In diesem Kapitel soll eine erste Begriffsbestimmung von Vertrauen im weitesten Sinne vorgenommen werden. Diese Begriffsbestimmung ist in einem zweifachen Sinne noch stark formal: Erstens gilt diese nicht dem Begriff interpersonellem Vertrauens im Besonderen, sondern zielt auf den allgemeinen Vertrauensbegriff ab. Vertrauen können wir schließlich nicht nur in Personen haben, sondern etwa auch in Institutionen, in Technik oder darauf, dass ein zukünftiges Ereignis geschehen wird. Den allgemeinen Vertrauensbegriff, der in all diesen Kontexten zur Anwendung kommt, werden wir im Folgenden auch den *generischen Vertrauensbegriff* nennen.

Zweitens beschränkt sich die Analyse darauf, zwei konstitutive Merkmale des Vertrauensbegriffes zu benennen. Oben haben wir zwar ausgeführt, inwiefern sich die in dieser Arbeit angestrebte philosophische Klärung des Begriffes interpersonellen Vertrauens nicht in der bloßen Angabe einer formalen Definition erschöpft, die lediglich verschiedene *definientia* des zu klärenden Terminus aufzählt. Schließlich zielt die letztendlich angestrebte Begriffsbestimmung darauf ab, eine umfassendere *Konzeption* interpersonellen Vertrauens zu erarbeiten. Jedoch gilt,

dass auch wenn das Aufzählen notwendiger Elemente oder *definientia* von Vertrauen dessen eigentliche philosophische Klärung nicht erschöpft, soll damit in diesem Kapitel ein erster formaler Umriss des Begriffes herausgestellt werden. Durch die Angabe der formalen Merkmale des Vertrauensbegriffes soll die begriffliche Grundstruktur abgesteckt werden, auf der die anschließende Untersuchung ihren Platz finden wird. Sie liefert damit eine begriffliche Grundorientierung.

Wie ich im Folgenden zeigen werde, ist Vertrauen im weitesten Sinne durch zwei Elemente charakterisiert: Erstens einem Risikobewusstsein und zweitens einer zuversichtlichen Einstellung. Zusammengenommen lässt sich Vertrauen somit als eine *zuversichtliche Haltung angesichts eines Risikos* begreifen. Bevor wir auf diese Analyse genauer eingehen, soll allerdings noch auf einen grundlegenden Vorbehalt eingegangen werden. So ist nämlich festzustellen, dass Vertrauen in Teilen der sozialwissenschaftlichen, aber auch philosophischen Literatur, in keinster Weise als eine Haltung aufgefasst wird. Vielmehr wird Vertrauen hier als das bloße *Eingehen* eines Risikos verstanden[16] – und zwar ungeachtet der genauen Einstellung, aus der heraus das Risiko eingegangen wird.

Ein Korollar dieser Sichtweise ist, dass sich Vertrauen unabhängig davon zeigt, ob man zuversichtlich ist oder nicht. Dieses Verständnis scheint immerhin auch in unserer Alltagssprache angedeutet, in der wir etwa so etwas sagen können wie: „Ich bin zwar pessimistisch, dass es gut ausgehen wird, aber ich muss mein Vertrauen in diese Person setzen“. Solche Redeweisen suggerieren, dass sich Vertrauen letztlich *allein* in unseren risikofreudigen Entscheidungen und Handlungen manifestiert – unabhängig von der Frage, ob man diese Entscheidung aus einer optimistischen oder pessimistischen Haltung heraus trifft. In der spieltheoretischen Terminologie könnte man sagen, dass sich Vertrauen allein darin zeige, dass bloße Vertrauens*akte* oder Vertrauens*handlungen* eingegangen werden, nicht in den Einstellungen, die diese womöglich unterlegen.[17]

Als Reaktion auf diese abweichende Begriffsbestimmung soll hier die These vorgeschlagen werden, dass das *Wort* „Vertrauen“ eine Ambiguität aufweist. Es kann auf zwei verschiedene *Begriffe* verweisen, die wir hier auseinanderhalten werden: Einerseits auf Vertrauen als dem bloßen *Eingehen eines Risikos bzw. von Vertrauensakten* und andererseits auf Vertrauen als *vertrauender*

16 So etwa in Teilen der Spieltheorie, siehe z.B.: Coleman 1994, 100. Darüber hinaus verstehen viele Autor:innen, die sich auf sogenannte therapeutische Vertrauensformen beziehen, Vertrauen in diesem Sinne. Therapeutisches Vertrauen bezeichnet dabei Vertrauensakte, die den Vertrauenspartner (kurz- oder langfristig) dazu bewegen sollen sich vertrauenswürdig zu verhalten. Klassisch hierzu: Horsborough 1960. Ausführlicher werden wir diesen Punkt im Kap. 7 diskutieren.

17 Vgl. Lahno 2002, 33.

oder vertrauensvoller Haltung. Demnach mag man zwar zuweilen unter „Vertrauen“ schlichtweg das Eingehen eines Risikos verstehen. Normalerweise scheinen wir mit dem Wort „Vertrauen“ aber auf eine spezifische Haltung zu referieren, typischerweise etwa dann, wenn wir sagen, dass wir Vertrauen (in etwas/jemanden) *haben* oder auch wenn wir sagen, dass wir *darauf* vertrauen, dass etwas passieren wird. In diesen Redeweisen verweist der Terminus „Vertrauen“ deutlich auf eine bestimmte Einstellung und keine bloße Handlungsdisposition. In der Regel verstehen wir Vertrauen also, wie Richard Holton es ausdrückt als „a distinctive kind of attitude involving a distinctive state of mind“.[18] Dies ist zugleich auch die Standardperspektive auf das Vertrauensphänomen in der Literatur. Entsprechend bewegen sich die eigentlichen philosophischen Debatten auch um die Frage, wie genau diese Haltung zu verstehen ist. Aus diesen Gründen werden wir uns im weiteren Verlauf der Arbeit, sofern nicht ausdrücklich anders gesagt, *ausschließlich auf Vertrauen als Haltung* beziehen. Diese zeigt sich zuweilen zwar in der Bereitschaft zum Eingehen eines Risikos, umgekehrt aber bedeutet die bloße Risikobereitschaft als solche noch nicht, dass man eine vertrauende Einstellung hat.

2.1.1 Vertrauen und Risikobewusstsein

Betrachten wir für eine erste Annäherung an die begriffliche Struktur von Vertrauen zwei Beispielfälle. Der erste Fall stellt ein Beispiel dar, in dem sich das Vertrauen auf eine Person bezieht, der zweite, in dem dies nicht der Fall ist. Dabei ist der erste Fall an ein berühmtes Beispiel von David Hume angelehnt.[19] Da wir wiederholt auf dieses Beispiel zurückkommen werden, geben wir ihm den Namen *Heu-Ernte-Beispiel*:

> Bauer A und Bauer B stehen beide kurz vor der Heuernte. Keiner von ihnen jedoch kann die Ernte je für sich alleine einfahren. Sie sind vielmehr auf wechselseitige Hilfe angewiesen. Stellen wir uns nun vor, dass die Ernte von B zuerst ansteht, dann ergibt sich für A folgendes

18 Holton 1994, 63.

19 Die originalen Zeilen, die auch als klassische Problembeschreibung des Vertrauensdilemmas in der Spieltheorie gesehen werden können, lauten: „Your corn is ripe to-day; mine will be so to-morrow. ’Tis profitable for us both, that I shou’d labour with you to-day, and that you shou’d aid me to-morrow. I have no kindness for you,and know you have as little for me. I will not, therefore, take any pains upon your account; and should I labour with you upon my own account, in expectation of a return, I know I shou’d be disappointed, and that I shou’d in vain depend upon your gratitude. Here then I leave you to labour alone: You treat me in the same manner. The seasons change; and both of us lose our harvests for want of mutual confidence and security“ (Hume 2007a, 334).

> Problem: Für den Fall, dass er B hilft, hat er keine Garantie, dass B ihm im Gegenzug auch bei seiner danach anstehenden Ernte helfen wird. Insofern muss A, wenn er B hilft, darauf *vertrauen*, dass B ihn ebenfalls unterstützen wird.

Das zweite Beispiel illustriert einen Fall nicht-personellen Vertrauens:

> Bevor sich Tim auf den Weg zu einem Rendez-vous macht, überlegt er, ob es sich lohnt einen Regenschirm mitzunehmen. Er wirft einen Blick aus dem Fenster und erblickt einen klaren, blauen Himmel ohne Anzeichen für aufkommenden Regen. Zwar könnte er vorsichtshalber noch den Wetterbericht abrufen, aber angesichts der aktuellen Wetterlage, entscheidet er sich schlichtweg darauf zu *vertrauen*, dass es trocken bleiben wird. Deshalb verlässt er ohne Regenschirm das Haus.

An diesen beiden Fällen lässt sich ein grundlegendes begriffliches Merkmal von Vertrauen ablesen: Vertrauen ist eine Haltung, die wesentlich angesichts eines Risikos eingenommen wird. Wenn A in Vorleistung geht, geht er damit ein gewisses Risiko ein, ebenso wie Tim, wenn er ohne Regenschirm das Haus verlässt. Im Lichte dieses Risikos ist es für beide überhaupt begrifflich möglich, dass sie eine vertrauende Einstellung einnehmen können. Gäbe es kein Risiko, wäre es sinnlos die Kategorie des Vertrauens überhaupt aufzurufen.[20]

Ein Risiko besteht dabei grundlegend dann, wenn zwei Bedingungen vorliegen: Erstens muss eine epistemische Unsicherheit[21] über bestimmte Ereignisse herrschen. Entsprechend kann man nur unter der Bedingung vertrauen, dass nicht sicher ist, ob das, worauf man vertraut, auch tatsächlich der Fall ist bzw. der Fall sein wird.[22] In den vorliegenden Beispielen ist dies jeweils der Fall: Bauer A weiß nicht, wie sich Bauer B verhalten wird, und Tim weiß nicht zu 100 Prozent, ob es nicht doch regnen wird. Zudem muss sich diese Unsicherheit zweitens auf ein Ereignis beziehen, das für den Vertrauenden in irgendeiner Form *relevant* ist. Das heißt ganz allgemein, dass die Unsicherheit zulässt, dass die Entscheidung zu vertrauen für den Vertrauensgeber einen evaluativ negativen Ausgang nehmen kann. Im ersten Beispiel etwa besteht dieser darin, dass

20 Dies muss natürlich nicht bedeuten, dass tatsächlich ein Risiko besteht. Die Kategorie des Vertrauens wird auch dann relevant, wenn der Vertrauende nur denkt, es bestünde ein solches.

21 Ich differenziere hier nicht, wie etwa in den Wirtschaftswissenschaften üblich, zwischen berechenbaren Risiko und Ungewissheit, wobei im ersten Fall die Eintrittswahrscheinlichkeiten abgeschätzt werden können, im letzteren Falle jedoch nicht. Vgl. Müller 1993.

22 Entgegen unseren Beispielen ist Vertrauen nicht immer notwendig zukunftsbezogen. Ein Gegenbeispiel etwa wäre ein Fall von Sprechervertrauen (solange der Sprecher keine Aussage über die Zukunft getroffen hat): Ich vertraue einem Sprecher, dass, das was er sagt, *jetzt gerade* der Wahrheit *entspricht*.

Bauer A umsonst Bauer B bei der Ernte geholfen haben kann. Im zweiten Fall besteht er darin, dass Tim ohne Schirm womöglich in den Regen gerät. Im Fall interpersonellen Vertrauens hängt diese Unsicherheit im Übrigen wesentlich damit zusammen, dass man nicht weiß, wie sich die Vertrauensperson verhalten wird. Handelt sie nicht vertrauenswürdig, geht dies zulasten des Vertrauensgebers. In diesem Sinne kann man auch sagen, dass sich der Vertrauensgeber gegenüber dem Vertrauensnehmer verletzlich macht.[23]

Analog gilt auch, dass es begrifflich nicht möglich ist, in Bezug auf etwas zu vertrauen, das mit keinem Risiko verbunden ist, unabhängig davon, ob die erstgenannte oder die zweitgenannte Bedingung nicht erfüllt ist. Ich kann beispielsweise eine Ahnung haben, dass sich zurzeit eine Maus im Kölner Dom aufhält. Solange mir dieser Umstand vollkommen gleichgültig ist, hätte ich damit aber keinen Anlass, darauf zu *vertrauen*, dass dies der Fall ist. In diesem Szenario nämlich droht kein möglicher Schaden und damit auch kein Risiko. Gleiches gilt, wenn ein zukünftiges Ereignis für mich zwar relevant ist, aber ich mir absolut sicher bin, dass es eintreten wird. Aus diesem Grund schließt sich Vertrauen begrifflich mit Wissen aus. Sofern ich etwa weiß, dass nächste Woche mein Geburtstag sein wird, muss und kann ich nicht darauf vertrauen. Hier besteht für mich nämlich gar keine Unsicherheit.

Vertrauen kann man deshalb also nur, insofern man sich diesbezüglich eines Risikos *bewusst* ist. Diese Redeweise kann vielleicht missverständlich sein. Damit soll nicht gesagt sein, dass Vertrauen notwendigerweise mit einem *ausdrücklichen* Risikobewusstsein einhergeht. Ganz im Gegenteil möchte man sagen, dass sich besonders festes Vertrauen zuweilen gerade dadurch auszeichnet, dass Risiken nicht nachdrücklich erwogen werden müssen. Der Punkt ist hier nur, dass dies nicht heißen kann, dass die Vertrauende *keinerlei Wissen* vom bestehenden Risiko haben kann. Dann nämlich könnte ihr Vertrauen nicht mehr von bloß mangelndem Wissen über die Risiken unterschieden werden.[24] In einem gewissen, weit verstandenen Sinne, muss der Vertrauenden also ein Risiko bewusst sein, wenn auch nicht zwangsläufig in der Form eines *expliziten* Bewusstseins. Zur besonderen Form des impliziten Vertrauens kommen wir im Kapitel 9 aber noch genauer zu sprechen.

Insofern unter einem Risiko nur die bloße Möglichkeit eines „Schadensfalles“[25] gegeben ist, besteht unmittelbar auch die Möglichkeit, dass dieser Schadensfall nicht eintritt. Vertrauen besteht

23 Siehe bereits Annette Baiers Charakterisierung von Vertrauen als „accepted vulnerability“ (Baier 1986, 235).
24 Vgl. Kaminski 2013, 171.
25 Der Begriff des „Schadens“ muss hier natürlich einigermaßen weit verstanden werden. Es muss sich dabei nicht notwendig, um dramatische Ereignisse handeln. Ich kann etwa auch darauf vertrauen, dass mir mein Nachbar einen Kuchen zum Geburtstag überreichen wird. Es wäre aber nicht dramatisch, wenn er dies nicht tun würde.

also auch angesichts der Möglichkeit eines „positiven“ Ausganges. Diesen kann man auch als den Vertrauensgegenstand bezeichnen. Es ist das, w*orauf* man vertraut. In den obigen Beispielen vertraut Tim etwa darauf, dass es nicht regnen wird, sowie Bauer A darauf vertraut, dass Bauer B ihm helfen wird, und er seine Ernte nicht verliert. Das Vorhandensein dieser prinzipiellen Möglichkeit impliziert zudem für den Fall des interpersonellen Vertrauens, dass der Vertrauenspartner *grundsätzlich* in der Lage dazu sein muss, sich vertrauenswürdig zu verhalten. Unter anderem bedeutet dies auch, dass er prinzipiell die geeigneten Kompetenzen[26] dafür haben muss, das Vertrauen erfüllen zu können.

Wie erwähnt besteht nun eine bloße Vertrauensentscheidung darin, ein solches Risiko einzugehen. Eine bloße Risikobereitschaft als solche macht jedoch noch kein Vertrauen aus. Man kann auch bereit sein gewisse Risiken einzugehen, ohne tatsächlich Vertrauen zu haben, dass die Sache gut ausgeht. Dies deutet darauf hin, dass sich Vertrauen noch durch ein weiteres Moment auszeichnet.

26 Christian Budnik (2016) argumentiert gegen die Annahme einer solchen Kompetenz als notwendiger Bedingung von Vertrauen; eine These, die er mit KA abkürzt. Sein Hauptargument versucht nachzuweisen, dass technische Kompetenzen als Minimalbedingung von einem Sich-Verlassen herhalten können. Daraus scheint mir aber umgekehrt nicht zu folgen, dass es nicht auch für Vertrauen eine *notwendige* Bedingung sein kann – zumal, wie ich in der folgenden Arbeit argumentieren werde, man Vertrauen besser als ein Sich-Verlassen-auf *sui generis* verstehen sollte. Abgegrenzt ist es lediglich zu einer reduzierten Form des *bloßen* Sich-Verlassen-aufs.
Die These KA lässt sich insgesamt so verstehen, dass zu vertrauen begrifflich bedeuten muss, von den formal notwendigen Kompetenzen der anderen Person auszugehen, sich vertrauensgemäß verhalten zu können. Immerhin wäre es schlichtweg unsinnig zu sagen: „Ich vertraue dir, dass du ϕ tun wirst, aber ich glaube, du bist überhaupt gar nicht in der Lage ϕ zu tun“. Zugleich ist damit noch nicht gesagt, dass Vertrauen allein in der Unterstellung dieser Disposition *aufgeht*. Budnik gesteht dies zu (vgl. ebd., 112f.), bringt aber dennoch ein „Unbehagen“ darin zum Ausdruck, KA „direkt in eine Analyse des Begriffes von Vertrauen und Vertrauenswürdigkeit zu übernehmen“ (ebd.: 113). Woher genau sich dieses Unbehagen erklärt, bleibt eher angedeutet. Das erwähnte Hauptargument, das er hierfür anbringt, scheint mir nicht geeignet. Meines Erachtens deutet Budnik aber subtil einen weiteren möglichen Grund an, ohne ihn weiter auszuführen. Genauer verstehe ich ihn so, als schrecke ihn der Umstand ab, dass die Angabe von KA als notwendiger Bedingung eine gewisse *Form der Analyse* des Vertrauensbegriffes nahelegt: nämlich als der Angabe einer *Summe* von untereinander *unabhängigen* notwendigen Bedingung. Im Kontrast dazu deutet er in einem Paragraphen vielmehr einen holistischen Analysezugang an. Für ihn sei Vertrauen wesentlich aus der Bezugnahme zweier Personen zueinander zu verstehen. KA aber werde „völlig ohne“ diese Bezugnahme aufeinander formuliert: „Ob mein Freund Auto fahren kann, meine Hausärztin den Überblick über Kopfschmerzmedikamente hat oder mein Nachbar über einen grünen Daumen verfugt, hat zunächst gar nichts *mit mir,* der vertrauenden oder potentiell vertrauenden Person, zu tun.“ (ebd., 113). Dieser Punkt lässt sich so verstehen, als müsse die Zuversicht des Vertrauenden in die Kompetenzen des Vertrauenspartners *immer schon im Kontext* ihres Verhältnisses zueinander verstanden werden. Begnügt sich die Begriffsanalyse in der bloßen Auflistung von notwendigen Bedingungen, wie KA, birgt dies die Gefahr, dass solche internen Zusammenhänge untereinander unbeleuchtet bleiben. Ich bin mir nicht ganz sicher, ob Budnik eine solche methodische Überlegung zum Ausdruck bringen wollte oder nicht. Falls ja, dann stimme ich dieser in der Sache zu. Siehe hierzu auch die methodischen Überlegungen im Kapitel 1.2.

2.1.2 Vertrauen als zuversichtliche Haltung

Über das Risikobewusstsein hinaus zeichnet sich eine vertrauende Einstellung wesentlich durch die Annahme aus, dass nicht der Risikofall, sondern die positive Ausgangsmöglichkeit eintreten wird. Wir können dies auch so ausdrücken, dass die Haltung der Vertrauenden eine *zuversichtliche oder optimistische* Haltung[27] darstellt. Dies bedeutet wiederum gerade nicht, dass die Vertrauende das Risiko schlichtweg vergisst. Zuversichtlich kann sie gerade nur *angesichts* eines möglichen Risikos sein. Ohne Risiko wäre Zuversicht weder notwendig noch begrifflich möglich.[28] In Bezug auf interpersonelles Vertrauen hält Annette Baier entsprechend fest: „One leaves others an opportunity to harm one when one trusts, and also shows one's *confidence* that they will not take it.“[29]

Dass die Haltung der Vertrauenden zuversichtlich ist, macht sie mit einigen anderen Einstellungen und Handlungen unvereinbar. In erster Linie kann die vertrauensvolle Einstellung nicht widerspruchsfrei zugleich als eine zweifelnde Haltung beschrieben werden. Man sieht dies unmittelbar an der Unstimmigkeit einer Aussage wie: „Ich vertraue darauf, dass mein Brief pünktlich ankommen wird, aber ich zweifle daran, ob es tatsächlich passieren wird“. Das soll nicht heißen, dass man sein Vertrauen nicht hin und wieder hinterfragen kann. Der Punkt ist vielmehr ein grundlegend begrifflicher, nämlich: Dass, *insofern* man sein Vertrauen hinterfragt, man gerade nicht vertraut. Ähnliches gilt für das Suchen nach Kontroll- oder Schutzmechanismen gegen das Risiko, das im direkten Widerspruch zu einer vertrauenden Einstellung steht. Jemand, der aktiv Vorkehrungen trifft, damit der potenzielle Schadensfall nicht eintritt, offenbart damit gerade keine zuversichtliche und damit auch keine vertrauende Haltung. Nehmen wir etwa in dem obigen Beispiel an, dass Tim vorsichtshalber doch seinen Schirm mitnehmen würde, so würden wir womöglich zweifeln, ob er tatsächlich darauf vertraut, dass es trocken bleibt. Würde er wirklich vertrauen, stellte dies für ihn einen Grund dar auf den Schirm zu verzichten.

Die Charakterisierung einer vertrauenden Haltung als zuversichtlich ist hier übrigens absichtlich noch in einem Punkte vage formuliert. Eine in der philosophischen Literatur diskutierte Frage diesbezüglich ist etwa, ob es sich bei dieser zuversichtlichen Einstellung um eine *Überzeugung* handeln muss oder nicht. Diese Diskussion wird zwischen zwei Positionen geführt, die üblicherweise als Kognitivismus und Non-Kognitivismus bezeichnet werden. Da wir an dieser

27 Insbesondere herausgestellt hat dies Karen Jones (1998).

28 Hartmann (2011, 85-99) betont, dass dieser Optimismus auch *wahrhaftig* sein muss. Dies zeigt sich wesentlich in der Praxis, also daran, ob die Vertrauende auch *tatsächlich* ein Risiko eingeht. In diesem Sinne bezeichnet Hartmann Vertrauen auch als eine praktische Einstellung.

29 Baier 1986, 235; Hervorhebung von mir, J.W.

Stelle aber zunächst eine möglichst unkontroverse Analyse des Vertrauensbegriffes als Grundlage für die folgende Untersuchung einführen wollen, klammere ich diese Frage hier aus. Wir kommen am Anfang des dritten Teils dieser Arbeit darauf zurück. Um bezüglich dieser Frage vorerst neutral zu bleiben, formuliere ich, dass Vertrauen darin besteht davon *auszugehen, anzunehmen oder zuversichtlich zu sein,* dass das, worauf man vertraut, auch geschehen wird.

2.2 Interpersonelles Vertrauen als Vertrauen *sui generis*

Insgesamt haben wir nun generisches Vertrauen als *zuversichtliche Einstellung angesichts eines Risikos* definiert. Diese kann sich zudem im Eingehen von Vertrauensakten ausdrücken, ist aber nicht auf das Eingehen eines Risikos reduzierbar. Wohlgemerkt ist diese Charakterisierung generisch zu verstehen: Sie legt eine allgemeine begriffliche Struktur dar, die sowohl in Bezug auf interpersonelles Vertrauen zutrifft, als auch in Bezug auf nicht-personales Vertrauen, etwa, wenn wir der Technik, regelmäßigen Abläufen oder dem politischen System vertrauen.

In dieser Arbeit ist es uns jedoch in erster Linie um die Klärung interpersonellen Vertrauens zu tun. Es geht wesentlich darum zu klären, was es heißt *jemandem* zu vertrauen. Diesbezüglich mag nun die Frage aufkommen, weshalb dies eigentlich einer gesonderten Untersuchung bedarf. Man könnte ja meinen, interpersonelles Vertrauen wäre strukturell vollkommen identisch mit dem generischen Vertrauensbegriff, es unterscheide sich lediglich anhand seines *Gegenstandsbereiches* insofern es auf andere Personen gerichtet ist. Mit der philosophischen Klärung des weiten Vertrauensbegriffes erschöpfe sich damit im Wesentlichen auch die Klärung des Begriffs des interpersonellen Vertrauens, dem quasi lediglich der indexikalische Hinweis hinzugefügt werden müsse, dass es sich auf die Gegenstandsgruppe der Personen bezieht. Tatsächlich jedoch werden wir in den nächsten Kapiteln sehen, dass interpersonelles Vertrauen zwar weiterhin die formalen Merkmale des generischen Vertrauensbegriffes teilt, es darüber hinaus aber *strukturell-begrifflich zu verstehende und nicht nur objektbezogene* spezifische Differenzen zu nicht-personalen Vertrauensformen aufweist. Interpersonelles Vertrauen ist in diesem Sinne eine Vertrauensform *sui generis*.[30]

Wir werden im Folgenden anfangen die spezifische Form interpersonellen Vertrauens im Zuge einer Diskussion der Rational Choice-Theorie bzw. Spieltheorie des Vertrauens

30 Den Fall des in*tra*personellen Vertrauens (Selbstvertrauen) werden wir in dieser Arbeit insgesamt nicht betrachten. Siehe hierzu etwa: Dormandy 2020; Jones 2012a; Zagzebski 2012, Kapitel 2.

herauszuarbeiten. Wie wir sehen werden, operiert diese allein mit dem generischen Vertrauensbegriff und verfehlt damit den spezifischen Sinn interpersonellen Vertrauens. Auch wenn sich insbesondere die Spieltheorie mit Szenarien beschäftigt, in denen sich das Vertrauen auf andere Personen bezieht, so bezieht sie sich damit nicht auf Personen *als Personen*. Vielmehr sind Personen in der Perspektive der Rational Choice-Theorie prinzipiell auch durch nicht-personale Entitäten, etwa komplex agierende Mechanismen ersetzbar. Dies hängt, wie wir gleich sehen werden, insbesondere damit zusammen, dass die Spieltheorie die entscheidende normative Dimension interpersonellen Vertrauens nicht zu berücksichtigen vermag.

Es sei noch einmal wiederholt, dass der gesamte folgende zweite Teil dieser Arbeit nach Möglichkeit in Abstraktion von den epistemischen Fragen um Vertrauen behandelt werden soll. Dazu gehört insbesondere die Frage nach dem Kognitivismus, aber auch nach der Frage der Angemessenheit bzw. der Rechtfertigung von Vertrauen. Diesen werden wir uns erst im dritten Teil dieser Arbeit widmen. Im Folgenden geht es zunächst einmal um die grundlegendere Frage danach, was interpersonelles Vertrauen eigentlich ist, insbesondere wie seine normativen Zusammenhänge zu verstehen sind.

II. Teil: Zur interpersonellen Grammatik von Vertrauen

3. Die Rational Choice-Theorie des Vertrauens

3.1. Ziel des Kapitels

Vor allem in der gegenwärtigen sozialwissenschaftlichen Vertrauenstheorie sind Rational Choice-Ansätze, über eine Vielzahl an Disziplinen hinweg, weit verbreitet.[31] In diesem Zusammenhang spielen insbesondere darauf aufbauende spieltheoretischen Ansätze eine prominente Rolle, die auch Eingang in philosophische Auseinandersetzung gefunden haben. Allein aufgrund dieser Vorrangstellung in der Vertrauensliteratur scheint es geboten, auf die Rational Choice-Theorie des Vertrauens einzugehen. Was uns im Folgenden dabei vorrangig interessieren wird, ist der dabei – zumeist nur implizit – präsupponierte Vertrauensbegriff, den wir grundlegend klären und einordnen werden. Das Ziel wird es letztendlich sein aufzuzeigen, dass die Rational Choice-Theorie mit einem Vertrauensbegriff operiert, der bestenfalls dem generischen Vertrauensbegriff, wie wir ihn oben eingeführt haben, entspricht. In der Theoriesprache der Rational Choice-Theorie ist es jedoch nicht möglich die Besonderheit *interpersonellen* Vertrauens auf den Begriff zu bringen.

Wie wir zeigen werden, liegt dies im Kern daran, dass in der Perspektive der Rational Choice-Theorie Vertrauen lediglich einen risikogebundenen Kooperationsakt darstellt, der aus einer probabilistisch angegebenen, epistemischen Erwartung über das Verhalten des Kooperationspartners resultiert. Sie ist damit im Einklang mit unserer Definition des generischen Vertrauens, abstrahiert aber von vornherein von Besonderheiten in der Relation der beiden Vertrauenspartner zueinander. In einem ersten Schritt werden wir diesbezüglich zwei Aspekte herausstellen, die schon in frühen Beiträgen des philosophischen Vertrauensdiskurs als Spezifikum interpersonellen Vertrauens ausgewiesen wurden, in der spieltheoretischen Betrachtung aber aus dem Blick geraten. So geht Vertrauen in eine Person, erstens, nicht nur mit einer epistemischen Zuversicht bezüglich dieser, sondern auch mit einer an diese gerichtete normative Erwartung einher. Zweitens, und eng mit dem ersten Punkt verknüpft, ist der Vertrauende nicht nur optimistisch in Bezug auf die Handlung des Vertrauenspartners, sondern auch darüber, aus welcher *Haltung* heraus er handeln wird.

Für unsere Untersuchung werden wir die spieltheoretische Diskussion natürlich nicht allumfassend behandeln können und beziehen uns deshalb exemplarisch auf den Ansatz von James

31 Diese Perspektive ist insbesondere in den Wirtschafts- und Politikwissenschaften vorherrschend, darüber hinaus aber auch in der Soziologie vertreten. Für einen Überblick, siehe: Cook/Santana 2018. In der philosophischen Literatur finden sich ebenfalls spieltheoretisch orientierte Ansätze, beispielsweise: Dasgupta 1988, Gambetta 1988, Hardin 2002.

Coleman, den dieser in seinem soziologischen Grundlagenwerk *The Foundations of Social Theory* (1990) entwickelt. Dieser hat explizit versucht, das Phänomen interpersonellen Vertrauens mit den Mitteln der Spieltheorie rationalitätstheoretisch zu rekonstruieren.

3.2 Eine kurze Einführung in die Rational Choice- und Spieltheorie des Vertrauens

Grundlegend ist der Rational Choice-Ansatz des Vertrauens auf Kooperationsprobleme in einer besonderen Klasse von Entscheidungssituationen bezogen. Die im Fokus stehenden Entscheidungssituation gestalten sich dabei im Kern stets so, dass ein Akteur, der Vertrauensgeber, vor der Wahl steht, ob er eine kooperative Vorleistung gegenüber dem Vertrauensnehmer eingeht, die insofern riskant ist, als nicht sicher ist, ob der Vertrauensnehmer sich ebenfalls dazu entscheiden wird zu kooperieren.[32]

James Coleman geht es in seinem Ansatz darum, mit Mitteln der Rational Choice-Theorie individuelle Entscheidungen für oder gegen einen solchen risikoreichen Kooperationsakt zu erklären. Dafür setzt er die Vertrauenssituation in einem ersten vorbereitenden Schritt mit einer bloßen Wettsituation gleich: Die Überlegungen, ob man jemanden vertrauen sollte oder nicht, seien „nothing more or less than the considerations a rational actor applies in deciding whether to place a bet“[33]. Das kennzeichnende Merkmal einer solchen Wettsituation besteht dabei darin, dass eine risikoreiche Handlungsoption besteht, also eine Handlungsoption, die mit einer gewissen Wahrscheinlichkeit einen verlustträchtigen, mit der Restwahrscheinlichkeit aber einen gewinnbringenden Ausgang nehmen kann.

Die grundlegende Idee sei hier kurz anhand des oben beschriebenen *Heu-Ernte-Beispiels* erläutert. Demnach kann gemäß dem Analogievorschlag von Coleman Bauer As Entscheidung in Vorleistung zu gehen, indem er Bauern B bei der Ernte hilft, als bloßes Eingehen einer Wette verstanden werden. Der „Wetteinsatz“ bestünde hierbei in der Arbeitskraft, die er investiert, um B bei der Ernte zu helfen. Der „Wettausgang“ hängt davon ab, ob Bauer B ihm wiederum hilft: Tut er dies, hätte A die „Wette gewonnen“, tut er dies nicht, hätte er die „Wette verloren“.

Auf dieser Idee aufbauend unternimmt Coleman eine rational choice-theoretische Modellierung der Überlegung des Akteurs A. Dessen Vertrauensentscheidung basiert demnach auf

32 Brugger et al 2013, 68.
33 Coleman 1990, 99.

einem Kalkül, das dem Prinzip der Nutzenmaximierung folgt.[34] Entsprechend sollte der Akteur sich dann entscheiden zu vertrauen, wenn der alles in allem zu erwartende Netto-Nutzenertrag dieser Entscheidung höher als Null liegt; also höher als in dem Fall, in dem man nicht vertraut bzw. auf die Wette verzichtet. Die dabei zu berücksichtigen Variablen sind: Der Wetteinsatz (L), der mit einem Vertrauensakt einhergeht; der potenzielle Netto-Gewinn (G) im Falle gelingenden Vertrauens; und die Gewinnwahrscheinlichkeit (p), die als Einschätzung der Vertrauenswürdigkeit aufzufassen ist. Auf eine Formel gebracht, ist demnach eine Vertrauensentscheidung dann für eine Akteurin rational, wenn gilt:

$$p \times G - (1 - p) \times L > 0 \qquad \text{bzw. wenn} \qquad p/(1 - p) > L/G$$ [35]

Der Fall kann übrigens zu einem für die *Spieltheorie* relevanten Fall erweitert werden, wenn noch ein paar zusätzliche Bedingungen der Entscheidungssituation eingeführt werden. Genau genommen braucht es dafür drei weiterer Annahmen, die eine spieltheoretische Betrachtung zulassen.[36] Diese sind erstens, dass auch der Vertrauenspartner, in dem Beispiel Bauer B, ein ebenso nutzenmaximierender Akteur ist, und dass zweitens beide Akteure gemeinsam wissen, dass sie ebensolche nutzenmaximierende Akteure sind. Drittens muss zudem angenommen werden, dass beide gemeinsam um die Struktur der Handlungssituation, also um die „Regeln des Spieles" wissen. Unter diesen Annahmen würde das *Heu-Ernte-Beispiel* ein sogenanntes Vertrauensspiel darstellen, das in der spieltheoretischen Literatur ausgiebig diskutiert wird, ohne dass ich an dieser Stelle auf diese Diskussion eingehen kann.[37]

Relevanter für uns sind die grundlegenden begrifflichen Konzeptionsentscheidungen des Ansatzes. Hierbei ist zunächst zu beachten, dass Coleman Vertrauen allein als eine Handlung,

34 Ebd.: 14, 96. Ebenso: Hardin 2002, Kap.1; Ripperger 1998, Kap. 4.

35 Entsprechend ist die Entscheidung irrational wenn gilt $p/(p-1) < L/G$ und indifferent im Falle $p/(p-1)=L/G$. Die zugrundeliegende Entscheidungsregel ist die Bayes-Formel oder Erwartungswertregel.

36 Vgl. die Darstellung der vier grundlegenden Annahmen der Spieltheorie in Hargreaves-Heap/Varoufakis 1995, Kap.1.

37 Siehe die frühe Version von: Camerer und Weigelt 1988. Ebenso: Berg et al. 1995. Siehe auch die philosophische Diskussion des Problems in: Faulkner 2017, Hollis 1998, Lahno 2002, insbes. Kap.1-4. Ich gehe an dieser Stelle nicht auf die, auch innerhalb der spieltheoretischen Literatur geführte Diskussion um das Vertrauensdilemma ein (gleichwohl ich diese im Abschnitt 4.1 anreißen werde). An dieser Stelle geht es mir vielmehr grundlegend darum nachzuweisen, dass der Vertrauensbegriff der Spieltheorie entscheidende Aspekte interpersonellen Vertrauens nicht einzufangen vermag.

nämlich dem Eingehen eines Vertrauensaktes, versteht[38] und von der Betrachtung von Vertrauen als einer vertrauensvollen *Haltung* absieht. Ein Verständnis von Vertrauen als einer bloßen Handlung haben aber nicht nur wir oben kritisiert, es wird auch innerhalb der rational choice-theoretischen Literatur selber zuweilen als problematisch angesehen. Der optimistische Aspekt im Vertrauen ginge damit nämlich verloren und Vertrauen wäre, wie etwa Morton Deutsch festhält, nicht mehr von einem bloßen *Gambling* unterscheidbar.[39] In der Regel entscheiden sich deshalb auch an der Rational Choice-Theorie anknüpfende Autor:innen dazu, Vertrauen als mit einer (ausreichend hohen) epistemischen Erwartung einhergehend zu modellieren.[40]

Für die Argumentation dieses Kapitels können wir diese Bedingung akzeptieren und davon ausgehen, dass mit ihr der zuversichtliche Charakter einer vertrauenden Haltung adäquat eingefangen wird. Damit hätten wir es hier insgesamt mit einem Vertrauensbegriff zu tun, der im Einklang mit unserem Begriff generischen Vertrauens ist. So handelte es sich beim Vertrauen im rational choice-theoretischen Sinne im Wesentlichen um eine epistemisch zuversichtliche Haltung angesichts eines Risikos, die sich darin zeigt, dass Vertrauensakte eingegangen werden. Es bleibt jedoch dabei, dass im Rahmen der Theoriesprache der Rational Choice-Theorie die Besonderheit interpersonellen Vertrauens nicht repräsentiert werden kann.

Bevor wir uns dieser Problematik zuwenden, werfen wir einen genaueren Blick auf den rationalitätstheoretischen Hintergrund der Rational Choice-Theorie, wie man sie bei Coleman exemplifiziert findet. Der folgende Exkurs dazu ist allerdings nicht notwendig, um dem anschließenden Argumentationsgang zu folgen. Es bleibt deshalb der Leserin frei, diesen zu überspringen. Er dient lediglich dem besseren Verständnis und der Einordnung des Rational Choice-Ansatzes, der auch, unabhängig von der Vertrauensthematik, kritisch betrachtet werden kann. Außerdem scheint zumindest der zweite unten erläuterte Kritikpunkt an Colemans Vertrauensverständnis indirekt mit den grundlegenden rationalitätstheoretischen Annahmen zusammenzuhängen.

38 So erläutert er den Begriff Vertrauen über die Erläuterung des „placement of trust“, was aber nur das bloße Eingehen der Kooperation bzw der „Wette“ beschreibt – nicht etwa die dahinter stehende Haltung des Vertrauenden (beispielsweise: Coleman 1990, 94).

39 Deutsch 1958, 266.

40 Deutsch 1958; 1962; Hardin 2002, 10; Ripperger 1998: 88-90. Für eine Übersicht möglicher Verständnisse des genauen Verhältnisses zwischen Vertrauen und einer probabilistischen Erwartung: Siehe auch Lahno 1997, 312. So könne Vertrauen entweder als die Erwartung selber, als eine Folge der Erwartung oder als eine Grundlage der Erwartung verstanden werden.

3.3 Exkurs: Der humeanische Kern der Rational Choice-Theorie

Als humeanisch verstehe ich die Rational Choice-Theorie insofern, als sie im Kern zwei David Hume zugeschriebene Thesen übernimmt, die ich die *Trennungsthese* und die *Trägheitsthese* nenne.[41] Die Trennungsthese besagt dabei, dass ein kategorialer Unterschied zwischen Leidenschaften (*passions*) auf der einen Seite und die dem Verstandesvermögen zugeordneten epistemischen Ideen (*ideas*) auf der anderen Seite bestehen. In Humes originalen Worten:

> „A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification. When I am angry, I am actually possest with the passion, and in that emotion have no more a reference to any other object, than when I am thirsty, or sick, or more than five feet high. 'Tis impossible, therefore, that this passion can be oppos'd by, or be contradictory to truth and reason; since this contradiction consists in the disagreement of ideas, consider'd as copies, with those objects, which they represent.“[42]

Hume geht demnach davon aus, dass der menschliche Geist zwei grundlegend voneinander zu differenzierende Elemente aufweist, zum einen rein konativ verstandene Leidenschaften und zum anderen rein kognitiv strukturierte Ideen. Zugleich liefert er auch ein Kriterium zur Unterscheidung derselben: Ideen haben demnach stets einen repräsentierenden Gehalt, wohingegen es sich bei Leidenschaften um nicht-repräsentationale mentale Zustände handelt. Auch wenn dieses Kriterium umstritten sein mag[43], stellt die damit vollzogene dichotomische Weichenstellung den Kern solcher Theorien praktischer Rationalität dar, die gemeinhin als humeanisch oder neo-humeanisch bezeichnet werden.

Ideen stehen nun aufgrund ihres propositionalen Gehaltes in inferentiellen Beziehungen

41 Chwaszcza 2003, §2. Ob Hume wirklich so eindeutig Humeaner im hier rekonstruierten Sinne war, kann allerdings bestritten werden (vgl. Milgram 1995). Mir geht es hier allerdings weniger um eine angemessene Exegese des Werkes Humes, sondern vielmehr um eine Diskussion einer einflussreichen Rekonstruktion seiner Theorie praktischer Rationalität.

42 Hume 2007a, 266f.

43 Hervorzuheben ist hierbei insbesondere der Vorschlag von Michael Smith in: Smith (1987). Smith reagiert damit kritisch auf Humes Tendenz zu einer „«strong phenonmenological conception» of desires“ (ebd, 45). Demnach verstehe Hume Wünsche (*desires)* wesentlich als unmittelbar bewussten Empfindungszuständen. Damit unterschlage er jedoch unter anderem deren propositionalen Gehalt. Ein Wunsch ist schließlich immer ein gehaltvoller Wunsch „ein bestimmtes ϕ zu tun“. Außerdem scheint es Fälle zu geben, in denen man einen Wunsch haben kann, ohne, dass sich dieser permanent in einer entsprechenden bewussten Empfindung niederschlagen muss. Smith nennt diesbezüglich etwa den Fall eines Vaters, der den generellen Wunsch hat, dass es seinen Kindern gut gehe (ebd.: 48). Als Gegenvorschlag, um angemessen zwischen *desires* und *beliefs* zu unterscheiden führt Smith deshalb ein *direction-of-fit*-Kriterium ein. Während mit Überzeugungen wesentlich der Anspruch verbunden ist, dass diese den Tatsachen der Welt entsprechen, gelte umgekehrt für jemanden, der einen Wunsch hat „that he is in a state *with which the world must fit“* (ebd.: 54). Zur Kritik an diesem Vorschlag: Chwaszcza 2003, §4.2.

zueinander. Aus Ideen können prinzipiell weitere Ideen logisch gefolgert werden. Insofern nach Hume Leidenschaften jedoch keinen propositionalen Gehalt aufweisen, kann man auf sie grundlegend nicht im Rahmen vernunftgeleiteter Überlegungen schließen. Dies hat das zweite berühmte Humesche Diktum von der *Trägheit der Vernunft* zur Folge. Geht man nämlich zusätzlich davon aus, dass Leidenschaften im Gegensatz zu Ideen auch Motivationen für Handlungen darstellen, dann folgt insgesamt, dass ideenbezogene Überlegungen, keinen Einfluss auf unser Handeln haben können:[44]

> „Since reason alone can never produce any action, or give rise to volition, I infer, that the same faculty is as incapable of preventing volition, or of disputing the preference with any passion or emotion.on alone cannot be a motive to any action of the will. […] Reason is and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them.“[45]

Diese radikal formulierte Schlussfolgerung gehen späteren Neo-Humeanern in der Regel zu weit. Sie gestehen vielmehr zu, dass praktische Überlegungen Einfluss auf unser Handeln haben können. Allerdings reichen diese Überlegungen nur soweit, als sie uns über die Bedingungen zur Realisierung unserer Leidenschaften aufklären können. Diese gemeinhin Instrumentalismus genannte Vorstellung, können wir im Folgenden genauer am Beispiel der neo-humeanischen Position von Bernard Williams erläutern.

In seinem klassischen Aufsatz *Internal and External Reasons (1979)* formuliert Williams eine Kritik an der sogenannten externalen Auffassung praktischer Gründe, nach der Gründe und Motivationen in *keinem* begrifflichen Zusammenhang stehen. Williams erläutert diese Vorstellung anhand der Figur Owen Wingrave aus Henry James‘ gleichnamiger Kurzgeschichte. In dieser Geschichte wird Owen durch seinen Vater dazu gedrängt der Armee beizutreten. Während Owen in keinster Weise dazu motiviert ist und er eine tiefe Abneigung gegen das Militär hegt, sieht sein Vater darin eine bedeutende und notwendige Pflichterfüllung. In diesem Zusammenhang mag Owens Vater nun sagen, dass sein Sohn einen Grund dazu hat in der Armee zu dienen. Diese Rede bezieht sich jedoch auf einen Grund im externalen Sinne, da dieser in keinster Weise für Owen selber motivational ausschlaggebend ist.

Williams moniert nun an dieser externalen Vorstellung von Gründen, dass solcherlei kein Potenzial zur Handlungserklärung mehr besäßen.[46] Zur Erklärung einer Handlung durch Gründe

44 Vgl. Koorsagard 1986.
45 Hume 2007a, 266.
46 Williams 1979, 22.

müsse deshalb auf ein motivationales Element des Akteurs Bezug genommen werden. Auf dieser Grundlage liefert uns Williams folgende alternative Definition praktischer Gründe: „[An actor] A has a reason to ϕ iff A has some desire the satisfaction of which […] A believes will be served by his ϕ-ing.“[47] Diese Idee illustriert er anhand des folgenden Beispieles: Angenommen eine Person möchte einen Gin-Tonic trinken und ist der Überzeugung, dass vor ihr ein Glas Gin und ein Glas Tonic stehen, deren Inhalte zusammengemengt einen Gin-Tonic ergäben. In diesem Fall hätte sie einen Grund dafür, die Inhalte beider Gläser zusammen zu schütten und das Ergebnis zu trinken.

Es ist zu betonen, dass Williams Begriff der Wünsche (*desires*), der gewissermaßen an den Begriff der Leidenschaft (*passions*) von Hume anschließt, recht weit zu verstehen ist.[48] Tendenziell gilt für den Neo-Humeanismus des 20. Jahrhunderts, dass er den Begriff von *desires* maximal formal auffasst.[49] Einer Person in diesem Sinne einen Wunsch ϕ zu tun zu attributieren, bedeutet lediglich sie als eine Person zu beschreiben, die (kurz- oder langfristig) darauf abzielt, ϕ zu realisieren. Beispielsweise können unter Wünschen auch langfristige Ziele einer Person verstanden werden, die mit keiner besonders intensiven Gefühlsqualität einhergehen. Wie Williams betont gehören zur Kategorie der Wünsche demnach auch „dispositions of evaluation, patterns of emotional reaction, personal loyalties and various projects […] embodying commitments.“[50]

Auch wenn Williams nun weiterhin kategorial zwischen *desires* und *beliefs* unterscheidet, lässt Williams Definition zu, dass Akteure praktisch wirksame Überlegungen anstellen können. So könnte beispielsweise die erwähnte Person fälschlicherweise glauben, dass sich vor ihr ein Glas Gin befände, während es sich in Wirklichkeit um Petroleum handele. Wir könnten sie auf diesen Fehler aufmerksam machen und sie somit von einer Handlung argumentativ abbringen: Ein Akteur kann in diesem Sinne „come to see that he has reason.“[51] Menschliches Handeln lässt sich also durch praktische Überlegungen anleiten, und zwar in der Form, dass in praktischen Überlegungen die bestmöglichen Mittel zur Befriedigung gegebener *desires* eruiert werden können. Obsolet ist damit in jedem Fall der mögliche Vorwurf gegenüber huemanischen Theorien, sie können

47 Ebd., 17.

48 Diese terminologischen Ausdrücke *belief* und *desire* waren dabei schon im Zusammenhang des sogenannten *desire-belief*-Modell üblich. Siehe zu diesem Modell den *locus classicus*: Davidson 1963.

49 Damit scheint Williams die Herausforderung umgehen zu können anzugeben, welche Wünsche es letztendlich sind, die rationales Handeln begründen. Im Einklang damit enthalten sich Neo-Humenaer typischerweise bezüglich substanzieller Aussagen über das Wesen von Wünschen und benutzen den Begriff in einem formalen Sinne. Statt von Wünschen spricht man dann auch von sogenannten *pro-Attitüden*. Diese Rückzugsoption findet man beispielsweise ausgeführt bei: Davidson (1980, 13) und Smith (1987, 55). Definiert werden können diese pro-Attitüden dann über eine formale Eigenschaft, etwa bei Smith als „psychological states with which the world must fit“ (ebd., 55).

50 Williams 1979, 20.

51 Ebd.: 20.

faktisches Handeln nur naturgesetzesgleich beschreiben, aber in keinerlei Weise als Ergebnis einer an einem praktischen Standard orientierten Überlegung charakterisieren.

Man mache sich jedoch klar, dass praktisches Überlegen gemäß des neo-humeanischen Theorierahmens, starken Limitierungen unterliegt. Zwei Aspekte können wir diesbezüglich hervorheben. Bedeutsam ist erstens, dass die individuellen *desires* als solche nicht als rational reflektierbar verstanden werden können. Formen der Kritik und Reflexion derselben sind demnach entweder gar nicht möglich, oder müssen sich selbst wiederum an einem höherstufigen *desire* orientieren.[52] Damit entscheidet sich insgesamt also die Rationalität einer Handlung letztlich an der Wahl der besten Mittel zur Befriedigung eines unabhängigen, vorgegebenen und nicht rationalen Überlegungen unterstehenden *desires*. Dies kann man auch als neo-humeanische *Asymmetrie-These* bezeichnen[53]: Demnach haben *desires* den Status *fundamentaler* Elemente im Rahmen praktischer Überlegungen.

Damit zusammenhängend kann zweitens hervorgehoben werden, dass es sich bei *desires* stets um *akteursrelative* konative Elemente handelt. Die Rationalität einer Handlung wird damit allein in Relation zu den für *eine bestimmte* Person bestehenden Konationen beurteilt. Praktische Überlegungen beziehen sich lediglich auf die Überzeugungen, und helfen insofern zu spezifizieren, welche Handlungen zur Realisierung welchen Wunsches geeignet sind.

Die Rationalitätskonzeption der Rational Choice-Theorie nun baut entschieden an diesem (neo-)humeanischen Verständnis praktischer Rationalität auf. Entscheidend für diese ist lediglich zusätzlich noch, dass die verschiedenen Wünsche in einer numerischen Ordnung zu betrachten sind. Dafür müssen diese grundlegende formale Konsistenz- und Kontinuitätsbedingungen erfüllen (Konnexivität und Transitivität), um metrisiert, also jeweils Zahlenwerten zugeordnet werden zu können. Unter diesen Voraussetzungen spricht man terminologisch nicht mehr von Wünschen bzw. *desires*, sondern von *Präferenzen.*[54] Auf dieser Grundlage kann durch die Rational Choice-Theorie systematisch und quantitativ die Rationalität von Entscheidungen bemessen werden.

Trotz dieser rational-choice theoretischen Weiterentwicklung der Humeschen

52 Damit gesteht Williams durchaus zu, dass die motivationalen Elemente nicht einfach unveränderbar vorgegeben sind. Im Rahmen praktischer Überlegungen können etwa bestimmte Wünsche präziser erkannt werden oder wegfallen: „the process of deliberation can have all sorts of effects on [the motivational set] S“ (ebd., 20). Wie insbesondere John McDowell herausstellt, kann der Prozess der Deliberation im Rahmen der internalistischen Theorie letztlich dennoch nur durch vorgegebene desires 'kontrolliert' gedacht werden (McDowell 1995, 71). Der Wunsch „zu trinken, was in dieser Flasche ist“ kann beispielsweise kritisiert werden, wenn man darauf aufmerksam macht, dass sich Petroleum in der Flasche befindet. Dies setzt dann aber gemäß der internalistischen Position voraus, dass die Person wiederum den Wunsch hat „kein Petroleum zu trinken“.

53 Vgl. Halbig 2003, 138.

54 Auf die genauen Details will ich an dieser Stelle nicht eingehen. Siehe hierzu genauer: Hahn 2017, 144-162.

Rationalitätskonzeption bleiben auch bei Coleman die entscheidenden Weichenstellungen desselben erhalten. Diese sind, noch einmal zusammengefasst:

1. Eine grundlegende ontologische Trennung zwischen Präferenzen (bei Coleman in G und L zusammengefasst) und Überzeugungen (bei Coleman bezüglich p).
2. Eine Asymmetrie zugunsten der Präferenzen bezüglich der Handlungserklärung und -rechtfertigung. Praktisches Überlegen bezieht sich entsprechend darauf, auf welchem Wege die individuelle Präferenzen am besten realisiert werden können.
3. Ein Verständnis von Präferenzen als rein akteursrelative psychologische Entitäten. Es bleibt bei dieser formalen Charakterisierung, die Möglichkeit von subjekttranszendierenden Bewertungen der Präferenzen als solchen werden nicht erwogen.

3.4 Kritik des Vertrauensbegriffes der Rational Choice-Theorie

Im Folgenden werden wir nun zwei spezifische Eigenschaften interpersonellen Vertrauens aufweisen, die in der rational choice-theoretischen Perspektive nicht eingefangen werden. Dabei bezieht sich die erste auf die normative Erwartung des Vertrauenden gegenüber der Vertrauensperson (1). Die zweite bezieht sich auf die vom Vertrauenden unterstellte Einstellung des Vertrauenspartners (2).

3.4.1 Die normative Erwartung des Vertrauenden

Schon in den Anfängen der aktuellen philosophischen Debatte über Vertrauen wurde die Idee formuliert, dass sich interpersonelles Vertrauen durch eine normative Erwartung auszeichnet. Den Hintergrund dafür bildet üblicherweise der Versuch ein Distinktionsmerkmal von Vertrauen zu identifizieren, mit dem dieses von einem sogenannten bloßen Sich-Verlassen abgegrenzt werden kann. Bloßes Sich-Verlassen ist dabei ein von Annette Baier[55] eingeführter Gegenbegriff zu Vertrauen, der lediglich die Haltung desjenigen beschreibt, der sich in seinen praktischen Planungen auf das vorhersehbare Verhalten anderer einstellt. Dies hat sie prominenterweise anhand

55 Erstmals in: Baier 1986.

des folgenden Beispieles illustriert:[56] Stellen wir uns vor, die Königsberger Nachbarin von Immanuel Kant nutzt die Tatsache, dass Kant regelmäßige Spaziergänge abhält, als eine Art Zeitangeber; nehmen wir etwa an, dass sie davon ausgeht, dass es Zeit ist, das Abendessen zuzubereiten, sobald Kant an ihrem Haus vorbeigeht. In diesem Fall würde sich die Nachbarin, so Baiers Formulierung, lediglich auf Kant *verlassen,* ihm aber nicht *vertrauen.* Obwohl Kants Nachbarin eine gewisse optimistische Perspektive bezüglich Kants Verhalten hat, und obwohl sie sich damit potenziell gegenüber Kants Verhalten verletzlich macht, könne man hier, so Baier, nicht davon sprechen, dass sie *Vertrauen in Kant* hätte.

Sich-Verlassen fällt damit zwar unter die Kategorie des Vertrauens im generischen Sinne. Was hieße es jedoch spezifischer *jemandem zu vertrauen*? Schon Baier versuchte in Bezug auf diese Frage das Spezifikum interpersonellen Vertrauens symptomatisch an den typischen Reaktionen im Fall einer Enttäuschung festzumachen. Wird nämlich eine Verlässlichkeitsannahme nicht erfüllt, kann man höchstens enttäuscht oder verärgert reagieren*,* so wie man sich etwa darüber ärgern würde, wenn man bei einer Wette verlöre. Das Vertrauen in jemanden zu verletzen, bedeute jedoch den Vertrauenden zu *verraten oder im Stich zu lassen*: „trusting can be betrayed, or at least let down, and not just disappointed."[57] Dies bedeutet, dass der enttäuschte Vertrauende andere Formen von Reaktionen zeigen kann: Er kann sich im Stich gelassen fühlen, empört sein oder dem Vertrauenspartner Vorwürfe machen – allesamt Reaktionen, die für die Nachbarin Kants in dem Beispiel schlichtweg abwegig wären.

Gemäß eines daran anknüpfenden Vorschlags von Richard Holton[58] lassen sich solche Arten von Reaktionen im Falle enttäuschten Vertrauens als sogenannte reaktive Einstellungen (*reactive attitudes*) im Sinne Peter Strawsons[59] auffassen. Damit sind allgemein gesagt solche Einstellungen und Gefühle gemeint, die wir gegenüber Personen zeigen, mit denen wir in einem Interaktionsverhältnis stehen, und mit denen wir auf deren wohlwollender, böswilliger oder indifferenter Einstellung uns selbst gegenüber respondieren. Die Liste der Beispiele, die Strawson nennt, umfasst unter anderem Dankbarkeit, Ärger, Vergebung, Liebe und verletzte Gefühle. Vertrauen zeige sich symptomatisch in der Möglichkeit einer solchen Art von Reaktion im Falle seiner Verletzung. Es zeichnet sich aus durch eine „readiness to feel betrayal should it be disappointed, and gratitude should it be upheld."[60] Vor diesem Hintergrund wird Vertrauen von

56 Ich beziehe mich ein wenig auf die Ausschmückungen des Beispiels von Bernd Lahno (2002).
57 Baier 1986, 235.
58 Holton 1994, 66.
59 Strawson 1968.
60 Holton 1994, 66.

Holton auch als eine teilnehmende Haltung (*participant stance*) tituliert.

Diese Erläuterungen der Begriffe der reaktiven Einstellung bzw. der teilnehmenden Haltung mögen eine Menge weiterer Fragen und Klärungsbedarf provozieren, die wir für unsere Zwecke jedoch ausklammern können. Uns geht es wesentlich darum, einen zentralen Aspekt der reaktiven Einstellungen hervorzuheben, der üblicherweise als der entscheidende identifiziert wird, um Vertrauen von Sich-Verlassen abzugrenzen: Demnach beruht eine solche reaktive Einstellung, wie das Gefühl sich betrogen zu fühlen, auf der *Verletzung einer normativen Erwartung*.[61] So hielt bereits Peter Strawson für reaktive Einstellungen im Allgemeinen fest: „[they] rest on, and reflect, an *expectation of*, and *demand for* the manifestation of a certain degree of goodwill or regard on the part of other human beings towards ourselves […].“[62] An unserem Beispiel illustriert hieße dies, dass wenn Kants Nachbarin Kant bezüglich seiner Spaziergänge *vertrauen* würde – ein Anlass könnte etwa sein, dass Kant ihr versprochen habe, immer zur gleichen Zeit an ihrem Haus vorbeizukommen[63] – sie nicht bloß eine epistemische Erwartung diesbezüglich, sondern auch eine entsprechende *normative* Erwartung Kant gegenüber hätte.[64] Würde Kant ihr Vertrauen verletzen, verletzte er damit diese normative Erwartung. Dies erklärt, weshalb sie in diesem Fall reaktive Einstellungen zeigen könnte.

Die normative Erwartung beim Vertrauen zeigt sich besonders prägnant daran, dass man sich nicht nur einfach im Stich gelassen *fühlen*, sondern dies der Vertrauensperson gegenüber auch in Form eines artikulierten *Vorwurfes* zum Ausdruck bringen kann.[65] Einen solchen Vorwurf können wir etwa mit Worten ausdrücken wie: „Du hast mein Vertrauen verletzt!“ oder „Du hast mein Vertrauen missbraucht!“ oder auch „Ich fühle mich von dir im Stich gelassen!“. Diese Vorwürfe lassen sich nicht einfach nur als *bloße* „Expression“ einer Gefühlslage verstehen. Sie bringen vielmehr die grundlegende normative Natur interpersonellen Vertrauens hervor. Sie bringen besonders deutlich zum Ausdruck, dass die andere Person, indem sie das Vertrauen

61 Vgl. die Interpretation Strawsons von Stephen Darwall (vgl. 2006, 74f.), Gary Watson (1987, insbes. 263f.) und Jay R. Wallace (1994): „there is an essential connection between the reactive attitudes and a distinctive form of evaluation […] that I refer to as holding a person to an expectation (or demand)” (ebd., 19).

62 Strawson 1968, 85; Hervorhebung von mir.

63 Die Frage, was angemessene Anlässe und Bedingungen dafür sind, werden wir unten ausgiebiger diskutieren (Kap.5.2/5.3).

64 In der Vertrauensliteratur wird dieser Punkt besonders eindrücklich von Margaret Urban Walker als Merkmal interpersonellen Vertrauens hervorgehoben: „In trusting one has *normative expectations* of others, expectations of others that they will do what they should and hence that we are entitled to hold them to it, if only in the form of rebuking and demanding feelings“ (Walker 2006, 80). Ohne auf Strawson zu referieren, sieht auch Philipp Nickel die normative Erwartung des Vertrauenden, genauer, die Unterstellung einer Pflicht (*obligation-ascription*), als das entscheidende Merkmal interpersonellen Vertrauens an (Nickel 2007).

65 So eindrücklich bei Hertzberg: “if I claim that A trusts B to do X, I commit myself to thinking that (in the absence of exonerating circumstances) B is open to blame if he fails to do it” (Hertzberg 1988, 319). Vgl. Lagerspetz 1998, 82.

verletzte, etwas tat, was sie nicht hätte tun *sollen.*

Prinzipiell gesagt kann gemäß diesem Verständnis eine vertrauende Haltung nur gegenüber *Personen* und nicht etwa gegenüber Artefakten oder Abläufen, eingenommen werden. Auch in diesem Folgepunkt unterscheidet sich interpersonelles Vertrauen vom bloßen Sich-Verlassen. Zwar kann man sich, wie in dem Beispiel, auch auf das Verhalten einer Person verlassen. In dieser Perspektive richtet man sich aber nicht auf die andere Person *als* eine andere Person. Sich auf x zu verlassen besteht schließlich, wie wir gesagt haben, lediglich darin, seinen praktischen Planungen die Überzeugung zugrundezulegen, dass x passieren wird.[66] In dieser Hinsicht werden auch Personen, auf die man sich verlässt, als bloßer Teil einer zu berechnenden Umwelt nivelliert. Die Nachbarin, die sich auf Kants Marotten verlässt, verlässt sich damit im Prinzip so auf Kant, wie man sich auch auf eine Taschenuhr verlassen könnte.

Wie gesagt: Die reaktiven Emotionen im Falle verletzten Vertrauens als einen Hinweis auf eine normative Erwartung im Falle verletzten Vertrauens zu sehen und diese damit als *differentia specifica* interpersonellen Vertrauens auszuweisen, stellt sicherlich eine Standardsichtweise dar, der wir uns hier anschließen werden. An dieser Stelle ist aber bereits schon einmal darauf hinzuweisen, dass wir diese Sichtweise im weiteren Verlauf der Arbeit entscheidend erweitern müssen. Wie wir im weiteren Verlauf der Arbeit zeigen werden, reicht es für eine umfassende Charakterisierung interpersonellen Vertrauens nicht, diese lediglich als mit einer normativen Erwartung einhergehend zu beschreiben.

Wenn wir uns für den Moment aber mit dieser Sichtweise zufrieden geben, sollten an dieser Stelle zwei grobe mögliche Missverständnisse diesbezüglich aus dem Weg geräumt werden: Erstens ist zu betonen, dass wir die Unterscheidung zwischen Vertrauen und *bloßem* Sich-Verlassen, aber nicht Sich-Verlassen *schlechthin* gezogen haben. Vielmehr scheint es sinnvoll Vertrauen als eine Form *sui generis* eines Sich-Verlassen zu verstehen. Wenn Person X Person Y bezüglich einer Handlung ϕ vertraut, heißt dies immerhin *auch*, dass X davon *ausgeht*, dass Y ϕ tun wird, und in seinen praktischen Planungen von ebendieser Annahme ausgehen kann. Dies ist es immerhin, was wir bei der Definition von Vertrauen im generischen Sinne als Zuversicht im Vertrauen beschrieben haben. Wenn beispielsweise die Nachbarin Kant *vertraut* sein Versprechen einzuhalten, dann *verlässt* sie sich auch darauf und würde etwa keinen Anlass sehen, sich eine Uhr

66 Auch diese Charakterisierung ist an Richard Holtons Definition angelehnt (1994, 3). Zu beachten ist aber, dass Holton Sich-Verlassen nicht als die Überzeugung, dass etwas passieren wird versteht, sondern vielmehr allgemeiner als eine praktisch wirksame Annahme, die etwa auch die Form einer Hypothese bestehen kann. In Übereinstimmung mit einem breit geteilten begrifflichen Verständnis werde ich Sich-Verlassen aber ausschließlich auf epistemischen Überzeugungen beruhend verstehen.

besorgen zu müssen oder Ähnliches.[67] Vertrauen bedeutet also immer *auch*, sich auf etwas zu verlassen, ist aber zugleich *mehr* als das.

Zweitens ist zu betonen, dass Normativität hier nicht mit Moralität gleichzusetzen ist. Der Umstand, dass Vertrauen mit einer normativen Erwartung einhergeht, sagt noch nichts über die letztendliche moralische Bewertung der Vertrauensbeziehung als solcher aus. Auch Bankräuber beispielsweise müssen sich untereinander vertrauen (gerade sie!), und dies schließt mit ein, dass sie auch normative Erwartungen an den jeweils anderen haben bzw. dem anderen Vorwürfe machen, wenn er das Vertrauen enttäuscht. Davon zu trennen ist aber die Frage nach der letztendlichen moralischen Bewertung dessen, *worauf* sie einander vertrauen. Hier kann man zumindest aus einer externen, kritisch bewertenden Perspektive zu dem Ergebnis kommen, dass der Vertrauensgegenstand unmoralischer Art ist. Diese Frage nach der moralischen Bewertung des Inhalts des Vertrauens ist aber eine andere Sache als die normative Erwartung aus der Perspektive des Vertrauenden an den Vertrauenspartner.[68]

In jedem Fall lässt sich mit den begrifflichen Mitteln der Rational Choice-Theorie eine solche normative Erwartung nicht darstellen. Der Begriff des Vertrauens, wie ihn die Rational Choice-Theorie, etwa bei Coleman, veranschlagt, kommt allenfalls einer Form von Sich-Verlassen gleich. So modellieren rational choice-theoretische Ansätze Vertrauen entweder in Form eines bloßen Aktes (so wie Coleman) oder maximal in Form epistemischer Erwartungen oder Überzeugungen, finden jedoch eine Grenze darin normative Zusammenhänge zu beschreiben. Entsprechend kommt die Vertrauenspartnerin auch nicht als eine Person in den Blick, sondern vielmehr als ein zu kalkulierender Umstand der Umwelt. Dabei hilft auch die *spieltheoretische* Erweiterung der Perspektive nicht weiter. Diese versteht die Vertrauenspartnerin zwar als eine *Akteurin* mit Präferenzen und Überzeugungen. Damit aber macht sie letztlich nur die Prognose ihrer Handlung komplexer (immerhin muss sie nun auch das nutzenmaximierende Kalkül des Anderen mit einbeziehen). Die Vertrauenspartnerin ist damit höchstens als eine *besonders komplexe* Variable im nutzenmaximierenden Kalkül anzusehen, aber immer noch eine bloß zu

67 Auch Baier selbst geht im Übrigen davon aus, dass Vertrauen eine besondere Form des Sich-Verlassen-auf darstellt. Es sei demnach eine Form des „reliance on [the trustee‘s] good will toward one“ (Baier 1986, 234). Dass diese Charakterisierung von Vertrauen jedoch problematisch ist, da sie auch eine instrumentalistische Haltung gegenüber dem Wohlwollen des Anderen bedeuten kann, wird in Kap. 6.1 gezeigt.

68 Ganz und gar unabhängig sind diese beiden Fragen allerdings auch nicht. Kann sich ein Bankräuber *sinnvoll* über den verdeckten Ermittler der Polizei beschweren, sein Vertrauen gebrochen zu haben? Es scheint mir, dass man zumindest im Normalfall und aus der Sicht der Vertrauenspartner eine geteilte Bewertung des eigenen Vertrauensgegenstandes als eines nicht durch und durch unmoralischen haben muss. Der Punkt bräuchte eine tiefergehende Erläuterung, die ich an dieser Stelle nicht geben kann. Für eine abweichende Sichtweise siehe Karen Jones‘ Vergleich mit dem Verhältnis zwischen freundschaftskonstitutiven und moralischen Normen (Jones 2017, 97f.)

berechnende Variable.[69] Auch aus spieltheoretischer Sicht kann sie also nicht *als Person* beschrieben werden, gegenüber der man normative Erwartungen haben kann.

3.4.2 Die Einstellung des Vertrauensnehmers

Der Umstand, dass Vertrauen wesentlich mit einer normativen Erwartung an den Vertrauenspartner einhergeht, schlägt sich auch in unserem Verständnis davon nieder, was es eigentlich genau heißt, Vertrauen zu enttäuschen und damit auch, was im Kern Vertrauenswürdigkeit ausmacht. Unsere obige Definition der reaktiven Emotionen von Strawson deutete diesen Punkt schon an, da wir erwähnten, dass diese wesentlich auf die spezifische böswillige oder benevolente *Einstellung* einer Person und nicht einfach nur deren Handlungen reagieren. Auf diesen Punkt wurde insbesondere in der moralpsychologischen Diskussion in Bezug auf Vorwürfe und Tadel mehrfach hingewiesen.[70] So beziehen sich Vorwürfe nämlich in erster Instanz nicht auf bloßes Handeln einer anderen Person. Ihr primärer Gegenstand ist vielmehr die dahinter stehende Einstellung. Wenn mein Freund etwa meinen Geburtstag vergisst, dann kann dies Ausdruck einer indifferenten oder böswilligen Haltung sein. Wäre dem so, dann kann dies ein Anlass sein, ihm Vorwürfe zu machen. Wenn er meinen Geburtstag jedoch aus nachvollziehbaren, entschuldigenden Gründen und ohne böse Absicht vergisst, ist die *gleiche Handlung* nicht tadelnswert, da in ihr keine *Einstellung* zum Ausdruck kommt, die verwerflich wäre. Vorwürfe und andere reaktive Einstellungen stellen also wesentlich ein Respons auf die Haltung einer anderen Person dar.

Diese Analyse der reaktiven Einstellung, insbesondere von Vorwürfen, wäre ein möglicher Weg um aufzuzeigen, dass sich Vertrauen in eine andere Person auch auf die spezifische Einstellung derselben bezieht. Unabhängig von diesem Argument aber wurde im Vertrauensdiskurs im Prinzip die gleiche Idee aus einer anderen Richtung kommend motiviert, und zwar ausgehend von der Frage, was es eigentlich heißt Vertrauen zu *enttäuschen.* Das Phänomen des enttäuschten Vertrauens verweist nämlich ebenfalls auf die Einstellungen des Vertrauenspartners und nicht auf auf seine bloßen Handlungen. Nehmen wir etwa den folgenden, von Paul Faulkner beschriebenen, Fall zur Illustration zur Hand:

69 Entsprechend bleibt etwa auch Russell Hardin (2002), der auf die besonderen Motive eines vertrauenswürdigen Vertrauenspartners abhebt, immer noch insofern dem Theorierahmen der Spieltheorie verhaftet, als er Vertrauen allein als eine epistemische, aber keine normative Erwartung versteht (insbesondere: ebd., 20f.).

70 Siehe beispielsweise Wallace 2019, 81-86.

> „Suppose we've arranged to meet at a place close to your work for an after work drink. And you don't show. Then just as I am about to give up you turn up, but only because this is your route home and you are passing by. When I discover this, the initial annoyance at your tardiness will be replaced by a stronger feeling of resentment".[71]

In diesem Fall also würde mein Freund oberflächlich besehen das tun, worauf ich vertraut habe. Dennoch kann man sagen, dass er mein Vertrauen enttäuscht hat.[72]

Ein ähnliches Beispiel aus Debatte um Sprechervertrauen weist in die gleiche Richtung. So weist Elizabeth Anscombe auf den Fall hin, in dem ein Sprecher seinen Zuhörer eigentlich täuschen will, zugleich aber das Gegenteil dessen glaubt, was der Wahrheit entspricht.[73] In diesem Fall nun mag der Sprecher etwas aussprechen, was tatsächlich der Wahrheit entspricht, und damit von außen gesehen genau das Gleiche tun, was ein wohlinformierter und vertrauenswürdiger Sprecher tun würde. Dennoch würden wir in diesem Fall wohl kaum sagen, dass der Sprecher das Vertrauen in den Zuhörer voll und ganz erfüllt hat – immerhin hatte er eigentlich die Absicht ihn zu täuschen.

Diese zwei Beispiele unterstützen damit ebenfalls die These, dass die normative Erwartung im Vertrauen bestimmte Einstellungen seitens des Vertrauensnehmers erforderlich machen. Diese Einstellungen scheinen zudem auch genau solche zu sein, die wir einer vertrauenswürdigen Person zuschreiben würden. Die genauen Zusammenhänge hierzu werden wir zwar erst unten in der Diskussion um Vertrauenswürdigkeit aufweisen. Für den Moment kann aber vorläufig angenommen werden, dass die normativ verlangten Einstellungen eben diejenigen sind, die Vertrauenswürdigkeit ausmachen. Mit dieser Idee scheint die Unterscheidung zwischen Vertrauen und bloßem Sich-Verlassen ein Pendant seitens des Vertrauensnehmers zu kennen: Es spiegelt sich hier wider als Unterscheidung zwischen einer Haltung genuiner *Vertrauenswürdigkeit* und einem bloß vertrauensgemäßen Verhalten oder bloßer *Verlässlichkeit*.

Dass die andere Person vertrauenswürdig ist, stellt nun aber nicht nur den Gegenstand der normativen Erwartung im Vertrauen, sondern auch einen Teil der zuversichtlichen Einstellung dar. Im Gegensatz zu einem bloßen Sich-Verlassen auf das Verhalten eines Anderen, ist Vertrauen in eine Person gewissermaßen immer *doppelt zuversichtlich*. Wenn man einer Person P vertraut ϕ zu tun, geht man nicht nur davon aus, dass P ϕ tun wird. Man geht auch davon aus, dass sie dies *aus einer vertrauenswürdigen Einstellung heraus* macht. Der Optimismus, der Vertrauen im

71 Faulkner 2014, 1979.
72 Siehe auch die ähnlichen Beispiele in: Hinchman 2017.
73 Anscombe 2008, 4.

generischen Sinne eignet, erstreckt sich somit im Falle interpersonellen Vertrauens auch auf die Haltung des Vertrauensnehmers.

Dadurch ergibt sich insgesamt, dass Vertrauen auf zwei verschiedene Arten nicht erfüllt werden kann. In dem von Faulkner angeführten Fall etwa wird zwar nicht die Annahme, dass die andere Person wie versprochen *handeln* wird, enttäuscht. Die Annahme jedoch, dass sie dies aus einer vertrauenswürdigen Einstellung heraus macht, wird nicht erfüllt. Und diese Enttäuschung kommt zugleich einer normativen Enttäuschung gleich.

Anders sieht ein Fall aus, in dem der Vertrauenspartner nicht die Erwartung in seine vertrauenswürdige Einstellung enttäuscht, aber dennoch nicht die Handlung ausführt, in Bezug auf die man ihr vertraut hat. Angenommen mein Freund hat auf dem Weg zum Café einen Unfall und muss sich ins Krankenhaus begeben. Nehmen wir etwa auch an, er ruft mich umgehend an, um mich darüber zu informieren. In diesem Fall hätte er zwar nicht das getan, worauf ich vertraut habe. Dennoch hätte er sich mir gegenüber vertrauenswürdig gezeigt und insofern nicht mein Vertrauen verletzt.

Damit haben wir noch keine präzise Charakterisierung der vertrauenswürdigen Haltung gegeben. Sie ist in jedem Fall zu verstehen als eine Haltung, die in bestimmter Weise mit der normativen Erwartung des Vertrauenden zusammenhängt. Sie genauer zu klären kann auch nur vor dem Hintergrund eines grundlegenderen Verständnisses dieser normativen Erwartung geschehen und bleibt deshalb eine Aufgabe für die kommenden Kapitel. Der entscheidende Punkt besteht an dieser Stelle lediglich darin aufzuzeigen, wie unzureichend der Ansatz der Rational Choice-Theorie zur Beschreibung der vom Vertrauenden angenommen Einstellung des Vertrauenspartners ist. In der rational choice-theoretischen Perspektive wird der Kooperationspartner lediglich bezüglich seines Verhaltens betrachtet, dass sich mithilfe seiner Präferenzen und Überzeugungen prognostizieren lässt.

Bündeln wir abschließend die zwei genannten Punkte noch einmal in einer Kritik am raional choice-theoretischen Vertrauensverständnis. Grundlegend gilt, dass sich für diese Vertrauen wesentlich in Form einer bloßen Entscheidung für eine risikoreiche Kooperationsvorleistung darstellt. Damit abstrahiert sie von der dahinter stehenden normativen Dimension, die sich allerdings für den Begriff des interpersonellen Vertrauens und der dem Vertrauenspartner zugeschriebenen Einstellung als wesentlich erwiesen haben. Diese beiden Punkte und auch deren systematischen Zusammenhang näher zu beleuchten, wird Aufgabe der folgenden Kapitel sein.

Es sei noch einmal betont, dass die Kritik am Vertrauensbegriff der Rational Choice-

Ansätze lediglich insofern als Kritik zu verstehen ist, als es die genauen Grenzen dieses Begriffes aufzeigt. Diese Kritik enthält sich insofern der Diskussion des explanatorischen Wertes der Rational Choice- bzw. Spieltheorie.[74] Unser Ziel war lediglich, *deutlich* zu machen, dass damit ein *reduzierter* Begriff interpersonellen Vertrauens verwendet wird.

3.5 Zwischenfazit: Interpersonelles Vertrauen als Vertrauen *sui generis*

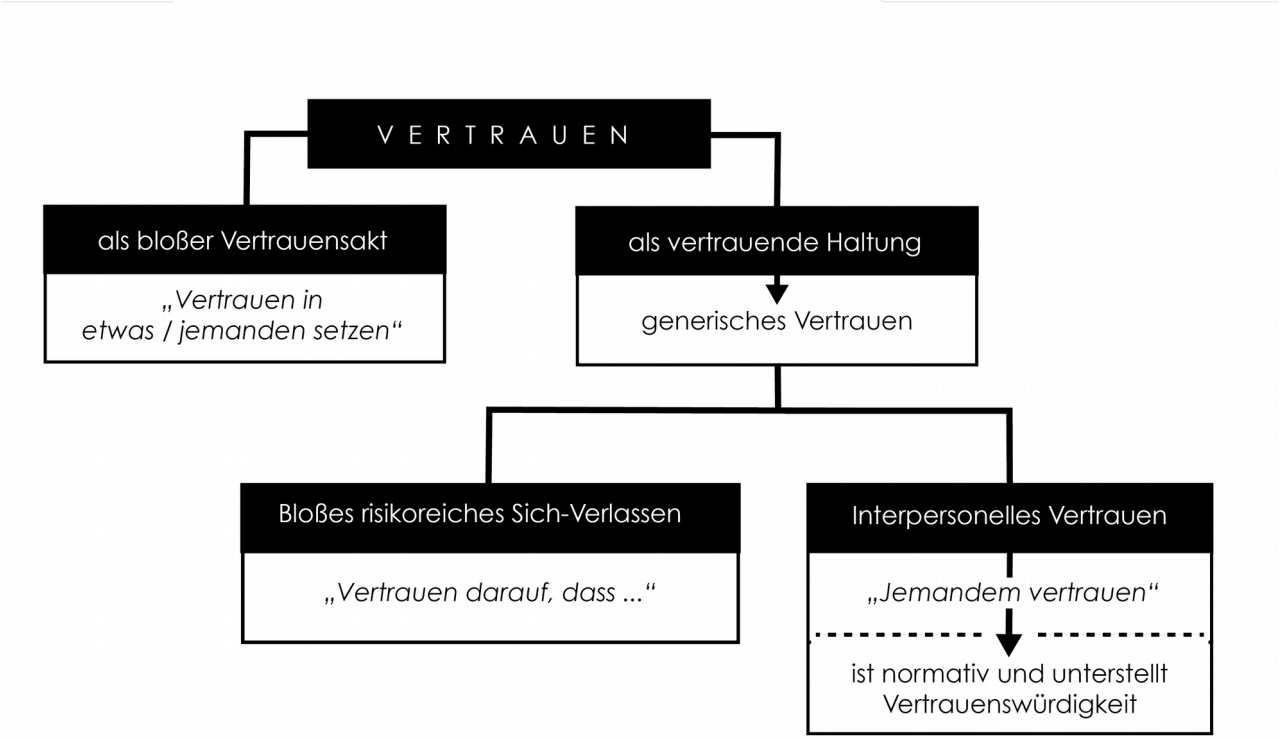
Wir sind nun in der Lage zu sehen, dass interpersonelles Vertrauen strukturelle Besonderheiten aufweist, die diese Vertrauensform als eine Vertrauensform *sui generis* ausweisen. *Interpersonelles Vertrauen* ist zwar weiterhin als Form generischen Vertrauens zu klassifizieren, insofern es mit einer zuversichtlichen Haltung angesichts eines Risikos einhergeht und sich in Vertrauensakten ausdrückt. Der Umstand aber, dass interpersonelles Vertrauen eine *normative* Erwartung beinhaltet und mit ihm dem Vertrauenspartner eine gewisse Einstellung attestiert wird, hebt diese Vertrauensform wesentlich ab. Dabei wird deutlich, dass das Spezifikum interpersonellen Vertrauens nicht einfach nur an seinem Gegenstandsbereich festzumachen ist. Im interpersonellen Vertrauen, so deutet sich hier an, wird vielmehr wesentlich eine andere Form der Einstellung dem Vertrauenspartner gegenüber eingenommen, der nun *als* Person – und nicht einfach ein zu prognostizierender Teil der Umwelt – vorgestellt wird.

Dieser Begriff interpersonellen Vertrauens ist zu kontrastieren mit einem bloßen risikoreichen Sich-Verlassen. Dieses stellt eine *zuversichtliche Einstellung* dar, die sich in der Bereitschaft dazu zeigt, Vertrauensakte einzugehen. Damit fällt bloßes Sich-Verlassen zwar auch unter den generischen Vertrauensbegriff. Im Gegensatz zum interpersonellen Vertrauen geht es allerdings *ohne* eine normative Erwartung und der Unterstellung einer vertrauenswürdigen Einstellung einher. Alltagssprachlich verwenden wir in diesem Zusammenhang übrigens zumeist auch die Vertrauensvokabel, in der Regel in Vertrauen-darauf-dass Konstruktionen. In dem Kant-Beispiel könnten wir etwa „Kants Nachbarin verlässt sich auf Kants regelmäßige Spaziergänge“ auch mehr oder weniger gleichbedeutend in die Aussage übersetzen: „Kants Nachbarin vertraut darauf, *dass* Kant seine regelmäßigen Spaziergänge abhält“.[75]

74 Philip J. Nickel (2017) unterscheidet in ähnlicher Weise von einer *restrictive* und *non-restricitve views of trust*. Dabei habe ein weites Verständnis des Begriffs von Vertrauen, wie es in der Rational Choice-Theorie vorherrscht, den explanatorischen Vorteil, dass mit ihm der Ursprung sozialer Kooperation und sozialer Institutionen besser erklärt werden könne.

75 Dass diese Formulierung nicht gleichbedeutend ist mit der Formulierung „Kants Nachbarin vertraut Kant, dass er

Insgesamt hat sich damit bisher die folgende Taxonomie verschiedener Vertrauensbegriffe ergeben, die hinter der alltagssprachlich verankerten Vokabel „Vertrauen“ stecken: Zum einen haben wir oben zwischen dem bloßen Eingehen eines Vertrauensaktes und einer vertrauenden Haltung im generischen Sinne unterschieden. Alltagssprachlich ist das bloße Eingehen eines Vertrauensaktes dabei mit Formulierungen wie „Vertrauen in jemanden/etwas setzen“ markiert. Eine vertrauende Haltung kann nun wiederum im Sinne des bloßen Sich-Verlassen oder im wahrhaftig interpersonellen Sinne verstanden werden. Erstere kann sich, wie eben gesagt, alltagssprachlich auch in „Vertrauen darauf, dass“-Formulierungen ausdrücken. Letzteres manifestiert sich in einer Formel wie „Jemandem vertrauen“ oder „Vertrauen in jemanden haben“. Graphisch können wir diese bisher getroffenen Unterscheidungen so zusammenfassen:



Graphik 1: Taxonomie verschiedener Vertrauensbegriffe hinter der alltagssprachlichen Vokabel „Vertrauen“.

Die folgenden Abschnitte dieses ersten Teils der Arbeit beziehen sich nun ausschließlich auf den Begriff des interpersonellen Vertrauens und versuchen diesen weitergehend zu klären. Zunächst geht es dabei um ein genaueres Verständnis der normativen Erwartung im Vertrauen und der Vertrauenswürdigkeit. In einem ersten Schritt werden wir uns dafür im Folgenden mit der

seine regelmäßigen Spaziergänge abhält“ zeigt ausführlich Benjamin McMyler (2011, Kap. 4.1).

Axiologie, das heißt des Zusammenhanges zwischen Vertrauen und Werten auseinandersetzen.

4. Die Axiologie interpersonellen Vertrauens

In der Auseinandersetzung mit dem rational choice-theoretischen Zugriff auf das Vertrauensthema konnten wir insgesamt zwei entscheidende Charakteristika interpersonellen Vertrauens herausarbeiten. Demnach umfasst dieses erstens nicht nur eine epistemische, sondern wesentlich auch eine normative Erwartung gegenüber der anderen Person. Diese normative Dimension konnten wir indirekt an den reaktiven Einstellungen im Falle enttäuschten Vertrauens aufzeigen. Zweitens beinhaltet Vertrauen nicht nur eine Erwartung über die Handlungen einer anderen Person, sondern auch darüber, aus welcher Einstellung heraus sie so handeln wird. Dabei ist die Art der Einstellung, die der Vertrauende dem Vertrauenspartner unterstellt, eine solche, die ihre Vertrauenswürdigkeit manifestiert.

Beide genannten Aspekte sollen im Laufe dieses und des folgenden Kapitels genauer verstanden werden. Die zwei entscheidenden Leitfragen für diese lauten also:

1) Wie genau ist die normative Erwartung des Vertrauensgebers gegenüber dem Vertrauensnehmer zu erklären?

2) Wie ist die Einstellung zu charakterisieren, die der Vertrauensgeber dem Vertrauensnehmer unterstellt? Was also ist das Wesen von Vertrauenswürdigkeit?

In diesem Kapitel wird es nun um die weitere Klärung dieser beiden Punkte gehen, indem wir ihren axiologischen Bezug herausstellen. Wenn Vertrauen mit einer normativen Erwartung einhergeht, scheint es naheliegend zu sein anzunehmen, dass sich diese auf einen bestimmten Wert oder Werte bezieht. Die normative Erwartung wird dann im Namen dieser Werte erhoben. Ebenso scheint sich eine solche axiologische Betrachtung zur Erläuterung des Begriffes der Vertrauenswürdigkeit anzubieten: Womöglich besteht die genuin vertrauenswürdige Einstellung darin, sich an diesem Wert oder diesen Werten zu orientieren. Die vertrauenswürdige Person lässt sich in ihrem Handeln also wesentlich von einem solchen Wert leiten. Gibt es hier also vielleicht schlichtweg einen solchen Wert/Werte, auf den sich dessen normative Erwartung und die

Einstellung des Vertrauenswürdigen beide beziehen?

Tatsächlich scheint es zunächst ein Allgemeinplatz zu sein, dass in gelingendem Vertrauen und damit auch in Vertrauenswürdigkeit als solcher ein Wert liegt. Worin genau ein solcher Wert bestehen sollte, wird in der Literatur zugegebenermaßen nicht immer klar herausgestellt. In jedem Fall scheinen damit oft unterschiedliche Sachverhalte gemeint zu sein. Zur zusätzlichen Verwirrung trägt dabei bei, dass zuweilen (ohne weitere Erläuterung) von einem *intrinsischen* Wert im Vertrauen und in Vertrauenswürdigkeit gesprochen wird. Das weitere Ziel dieses Kapitels ist es entsprechend auch, die Literaturlage diesbezüglich ein wenig zu sortieren. Dabei soll insbesondere der These vom intrinsischen Wert ein systematisch geklärter Sinn gegeben werden.

Bei der weiteren Klärung werden wir zunächst die These herausgreifen, dass Vertrauenswürdigkeit ein Wert zugeschrieben werden kann, der damit zusammenhängt, dass diese zu gelingenden Kooperationsbeziehungen beiträgt. Daran anschließend werden wir uns insbesondere auf die Frage fokussieren, wie genau der Zusammenhang zwischen einem solchen Wert, normativen Erwartungen und praktischen Gründen der Akteure zu denken ist. Dabei wird uns eine bestimmte Fehlauffassung, die wir mit Thomas Scanlon „teleologische Werttheorie“ nennen, als entscheidender instruktiver Gegner dienen. Die teleologische Werttheorie führt vor, wie sich die *praktische Orientierung* der beiden Vertrauenspartner am vertrauensbezogenen Wert gerade *nicht* verstehen lassen kann, da sie ein phänomenologisch unplausibles Bild der normativen Struktur zwischen den Vertrauenspartnern zeichnet. Insgesamt wird uns diese Diskussion die *besondere Art der Relation* der beiden Vertrauenspartner vor Augen führen, auf die wir im darauffolgenden Kapitel genauer eingehen werden.

Im Einzelnen werden wir in diesem Kapitel wie folgt vorgehen: Zunächst werden wir unabhängig von der hier zentralen Fragestellung auf die axiologischen Zusammenhänge in Bezug auf Vertrauen eingehen und klären, in welchem Sinne Vertrauen einen instrumentellen bzw. intrinsischen Wert hat (4.1). Anschließend widmen wir uns den beiden Leitfragen dieses Kapitels. Wir werden herausstellen, dass die axiologische Dimension von Vertrauen erstens die normative Erwartung des Vertrauenden begründen kann, und zweitens eine Erläuterung dafür liefert, was eine vertrauenswürdige Einstellung bedeutet (4.2). Letzterer Punkt beruht auf der Annahme eines allgemeinen Zusammenhanges zwischen praktischen Gründen und Werten. Damit wenden wir uns auch von dem (neo-)humeanischen Rationalitätsparadigma ab und einem wertbezogenen, man kann auch sagen aristotelischen, Paradigma praktischer Rationalität zu (Exkurs). Anschließend jedoch werden wir aufzeigen, dass der *genaue* Zusammenhang zwischen der Werthaftigkeit von

Vertrauen, der normativen Erwartung und der praktischen Einstellung des Vertrauenspartners nicht trivial, sondern erläuterungsbedürftig ist. Zu diesem Zwecke diskutieren wir die sogenannte teleologische Werttheorie und zeigen auf, wie diese im Bezug auf Vertrauen ein unplausibles Bild dieser Zusammenhänge liefert (4.4/4.5). Im Zuge dessen wird sich andeuten, dass sich die normative Dimension von Vertrauen wesentlich von der spezifischen Relation der beiden Vertrauenspartner her verstehen lassen muss, was zum nächsten Kapitel überleiten wird.

4.1 Der Wert in Vertrauen und Vertrauenswürdigkeit: Ein Sortierungsversuch

Dass Vertrauen und Vertrauenswürdigkeit in irgendeiner Form einen Wert[76] haben stellt einen weitgehend geteilten Allgemeinplatz dar. Unübersichtlicher wird es jedoch wenn man sich die Frage vorlegt, was dies eigentlich genau bedeutet. Mit der Rede von einem Wert scheinen hier nämlich zuweilen ganz unterschiedliche Sachverhalte gemeint zu sein, was sich zum Teil auch aus dem disziplinären Kontext ergibt. Zusätzliche Verwirrung kann dadurch entstehen, dass vereinzelt auch von einem *intrinsischen* Wert im Vertrauen und in Vertrauenswürdigkeit gesprochen wird, ohne, dass sich dieser Redeweise eine präzise Erläuterung anschließt. Im Folgenden werden wir deshalb ein wenig Sortierungsarbeit in Bezug auf die unterschiedlichen in der Literaturlandschaft vorfindlichen Bedeutungen und Begründungen der These von einem Wert bzw. intrinsischen Wert bezüglich Vertrauen und Vertrauenswürdigkeit leisten.

In einem ersten Schritt ist festzustellen, dass sich die Rede von einem „Wert des Vertrauen“ überwiegend auf dessen funktionalen Wert für intakte Kooperationsverhältnisse bezieht.[77] Hierbei kann Kooperation allgemein definiert werden als „das strukturierte, d.h. nicht zufällige, auf relative Dauer gestellte und sich durch (zumindest temporäre) wechselseitige Erwartungssicherheit (entweder auf reziproker oder nicht-reziproker Basis) begründende koordinierte (und damit zielorientierte) Wirken (Zusammenwirken) Mehrerer“.[78] Der Wert von Vertrauen und Vertrauenswürdigkeit ist somit durch ihren Status als kooperationsfördernde Ressourcen

76 Ich benutze hier den Terminus „Wert“ synonym mit „Gut“. Damit folge ich Jospeh Raz: „The notion of value used here, and in many philosophical discussions of axiological issues, does duty for a whole range of English words and expressions: 'value', 'good', 'best', etc., are just some of them. Others include many (but not all) uses of 'important', 'worthy', 'significant', 'meaningful', 'useful', 'meritorious', 'attractive', 'estimable', 'a good thing to do', and others“ (Raz 1999, 29).

77 Natürlich kann Vertrauen auch in anderen funktionalen Hinsichten wertvoll sein, etwa in seiner Rolle für die Persönlichkeitsentwicklung. Sein Beitrag zur Kooperation scheint aber der weitaus am verbreitetsten hervorgehobene zu sein.

78 Endreß 2012, 87.

begründet: Vertrauenswürdigkeit ist zum einen die entscheidende Bedingung dafür, dass die Kooperation überhaupt gelingt und damit auch, dass sie wiederholt wird. Vertrauen wiederum kann die initiale Bereitschaft und Vereinfachung zur Kooperation, auch unter unsicheren Bedingungen, bedeuten.[79] Nach Luhman etwa geschieht dies über eine Komplexitätsreduktion[80] und damit kognitive Entlastung durch Vertrauen, die eine entscheidende Bedingung zur Herausbildung und Stabilisierung kooperativer Interaktionsformen darstellt. Dieser kooperationsfördernde Wert von Vertrauen und Vertrauenswürdigkeit hat damit den Status eines funktionalen oder instrumentellen Wertes. Damit ist gemeint, dass sie allein insofern wertvoll sind, als sie als Mittel oder Bedingung zum Erreichen eines externen Zweckes, in diesem Fall gelingender Kooperation, dienen.

In dieser funktionalen Betrachtungsweise wird der Wert stabiler kooperativer Beziehungen wiederum in den verschiedenen Gesellschafts- und Sozialwissenschaften je unterschiedlich akzentuiert. Der Wert kooperativer Ordnungen wird so entweder als Grundlage des friedlichen sozialen Zusammenlebens, in Form von Sozialkapital als eine Bedingung funktionierender Demokratien[81] oder, ökonomisch betrachtet, als Mittel zur Senkung von Transaktionskosten dargelegt.[82] Zuweilen wird Vertrauen in seiner kooperationsermöglichenden Rolle auch ein *konstitutiv notwendiger Status* für das gesellschaftliche Zusammenleben zugeschrieben. In dieser Hinsicht werden funktionierende Vertrauensbeziehungen nicht zu Unrecht auch als „Grundlage des sozialen Zusammenhaltes“[83] bezeichnet, und Luhmann etwa hält prägnant fest, dass „Chaos und lähmende Angst“ die Alternative zu Vertrauen[84] wären. Dieser konstitutive Notwendigkeitscharakter von Vertrauen zeigt sich gerade für moderne, ausdifferenzierte, arbeitsteilige und technisierte Gesellschaften.[85] Arbeitsteilung bedeutet schließlich, dass Personen wechselseitig voneinander abhängig sind, also wechselseitig in Bezug auf das Verhalten anderer einem Risiko ausgesetzt sind. Durch die Professionalisierung und Technisierung der Positionen und Institutionen in der Gesellschaft ist es zudem zunehmend anspruchsvoller deren Arbeitsweise

79 Dies heißt nicht, dass es keine Kooperationsbereitschaft ohne Vertrauen geben kann: Ich kann etwa auch ein Versprechen akzeptieren, ohne der Person zu vertrauen, dass sie dieses einhalten wird. Vertrauen *erleichtert* aber die Akzeptanz eines Versprechens. Insofern stellt Vertrauen zumindest *typischerweise* eine Voraussetzung für Kooperation dar (Vgl. Cook et al. 2005).

80 Luhmann 2014, z.B. 28. Komplexitätsreduktion bedeutet dabei im Wesentlichen das Ausblenden von bestimmten Möglichkeiten in der Zukunft (ebd., 18, 26). Dies erinnert an unser obiges *definiens,* den Optimismus im Vertrauen: Im Vertrauen geht man von einem bestimmten positiven Verlauf der Ereignisse aus. Aus diesem begrifflichen Grund stehen Zweifel und Kontrollversuche im Widerspruch zu echtem Vertrauen. Statt eines begrifflichen Punktes, betrachtet Luhman dies vorrangig in Bezug auf seinen funktionellen Stellenwert, der in erster Linie darin besteht, dass Vertrauen kognitiv entlastend ist.

81 Klassischerweise etwa: Putnam 1992. Siehe auch: Sullivan/Transue 1999.

82 Siehe etwa Fukuyama (1995).

83 So der Buchtitel von Hartmann/Offe 2001.

84 Luhmann 2014, 1.

85 Vgl. Endreß 2002, 7.

nachzuvollziehen und unmittelbar zu kontrollieren. Vertrauen ist hier irreduzibel notwendig.

Aufgrund ihres axiologischen Zusammenhanges geht die Rede vom Wert von funktionierenden Kooperationen, von Vertrauen und von Vertrauenswürdigkeit zuweilen auch durcheinander.[86] Es ist entscheidend zu sehen, dass in der hier dargestellten funktionalen Perspektive auf Vertrauen der Wert von Kooperationen an sich axiologisch primär ist. Der Wert von Vertrauen und Vertrauenswürdigkeit besteht demnach nicht an sich, sondern nur *insofern* diese eine Bedingung gelingender Kooperation darstellen. Sowohl Vertrauen als auch Vertrauenswürdigkeit sind also nicht einfach bedingungslos wertvoll.

Über das axiologische *Bedingungsgefüge* in diesem Zusammenhang ließe sich noch einiges mehr sagen. Wir begnügen uns hier damit diesbezüglich nur zwei naheliegende Punkte hervorzuheben[87]: Erstens kann der kooperationsfördernde Wert von Vertrauen im konkreten Fall nicht vollkommen unabhängig von der Vertrauenswürdigkeit des Vertrauenspartners festgestellt werden. Naives Vertrauen, das in keinster Weise um die Vertrauenswürdigkeit des Interaktionspartners besorgt ist, kann auch vertrauensunwürdigem Verhalten den Vorschub geben und somit gerade funktionierende Kooperation verhindern oder unterminieren.[88] Nicht Vertrauen *per se*, sondern *bedachtes* Vertrauen stellt in diesem Sinne ein Gut dar. Zweitens kann unter bestimmten Bedingungen der Wert der Kooperation um die es geht selber in Frage stehen. Oben sprachen wir bereits einmal davon, dass Vertrauen auch moralisch fragwürdigen kooperativen Unterfangen dienen kann. Unter diesem Vorzeichen kann der Kooperation und damit auch dem darin situierten Vertrauen, bzw. der Vertrauenswürdigkeit in kollektiver Perspektive keinesfalls ein Wert zugeschrieben werden.

Neben diesem zentralem instrumentellen oder funktionalem Aspekt der Rede von einem Wert besteht auch eine zuweilen unklare Rede von einem intrinsischen Wert von Vertrauen und Vertrauenswürdigkeit – oder etwas, was dieser Sentenz im Wortlaut nahe kommt, wie etwa, dass es einen „inneren Wert“ im Vertrauen gäbe. Diese These findet man etwa in Bernard Williams

86 Siehe beispielhaft die an die Beiträge von Bernard Williams angelehnte Diskussion von Paul Faulkner, die dem Titel nach verspricht, sich mit dem Wert des Vertrauens (*The value of trust*) auseinanderzusetzen, letztlich aber den Wert von Ernsthaftigkeit und Genauigkeit (*sincerity and accuracy*) als spezifischen Ausprägungen von Vertrauenswürdigkeit herausstellt (Faulkner 2011, 176-179).

87 Bei genauerer Betrachtung können die Bedingungszusammenhänge im Einzelnen noch komplizierter sein. So denke man etwa an den von Piotr Sztompka als Demokratieparadox bezeichneten Umstand, dass institutionalisiertes Misstrauen langfristig gerade eine bessere Vertrauenskultur (und somit stabilere Kooperationsbedingungen) hervorbringen kann (Sztompka 1998). Man denke außerdem daran, dass Vertrauen auch strategisch als Manipulationsmittel dienen kann. So spezifiziert würde Vertrauen gerade Misstrauen säen und Kooperationsmöglichkeiten vereiteln.

88 So merkt etwa Russel Hardin an: „without [trustworthiness], there is no value in trust“ (Hardin 1996, 28-9).

einflussreicher Monographie *Truth and Truthfulness*[89] formuliert, in der der Autor von einem intrinsischen Wert von Vertrauenswürdigkeit spricht. Die genaue Bedeutung des Attributes „intrinsisch“ ist in diesem Zusammenhang jedoch zu erläutern. Dafür werde ich ein wenig ausholen müssen.

In seinen Überlegungen bezieht sich Williams vorrangig auf Sprechervertrauen, weshalb er auch primär von der Wahrhaftigkeit (*truthfulness*) oder auch Tugenden der Wahrheit (*virtues of truth*) als einer besonderen Form von Vertrauenswürdigkeit spricht. Wahrhaftigkeit setzt sich dabei wiederum aus der Ernsthaftigkeit (*sincerity*) und der Genauigkeit (*accuracy*) des Sprechers zusammen. Für diese nun stellt Williams zunächst in einem ersten Schritt ihren funktionalen Wert heraus. Mithilfe der Betrachtung eines vereinfachten fiktiven Naturzustandes veranschaulicht Williams die Bedeutung, die eine funktionierende Praxis der Zeugenschaft als eine bestimmte kooperative Praxis hat. Der Williamsche Naturzustand ist ein vereinfachtes Modell, um die Bedeutung funktionierender Zeugenschaftspraktiken im Kontext einer epistemischen Arbeitsteilung aufzuzeigen: In einem Zustand, in dem verschiedene Personen über verschiedene Weltausschnitte Wissen haben, ist es essentiell, dass sie dieses Wissen untereinander verlässlich *teilen* können.

In einem weiteren Schritt zeigt Williams auf, dass eine solche Zeugenschaftspraxis nur stabil sein kann, wenn Vertrauenswürdigkeit (bei ihm: *truthfulness*) in ausreichendem Maße ein intrinsischer Wert zugeschrieben wird. Um diese Idee zu motivieren bezieht er sich auch auf das Vertrauensdilemma in der Spieltheorie, weshalb ich auf dieses kurz eingehen werde. Als paradigmatisches Szenario zur Erläuterung desselben kann dabei wiederum unser Hume‘sches *Heu-Ernte-Beispiel* dienen. Das Dilemma kommt hierbei zustande, wenn man annimmt, dass man es, wie in der Spieltheorie vorausgesetzt, mit nutzenmaximierenden Akteurinnen zu tun hat, die zudem ebendies gemeinsam voneinander wissen. Unter diesen Bedingungen kann Bauer A damit rechnen, dass wenn er Bauer B bei der Ernte hilft, sich dieser aus Eigennutz nicht ohne Weiteres revanchieren wird. Für ihn hat es schlichtweg keinen Nutzen Bauern A im Gegenzug bei der Ernte zu helfen. Insofern Bauer A dies jedoch antizipieren kann, wird er von Anfang an keine kooperative Vorleistung gegenüber Bauern B eingehen. Somit wird die Möglichkeit zur Kooperation zwischen beiden im Keim erstickt. Damit jedoch können beide Bauern ihre Ernte nicht einholen, und sind dazu verdammt ein Ergebnis einzufahren, dass *beide* schlechter stellt, als wenn sie miteinander kooperieren würden. In der Sprache der Entscheidungstheorie verfehlen

89 Williams 2002.

somit die beiden Akteure das *pareto-optimale* Ergebnis.

Aus dieser Schlussfolgerung ergibt sich die Frage, wie es überhaupt möglich sein kann, dass wir uns alltäglich größtenteils problemlos vertrauen, und zwar ohne unterstellen zu müssen, dass wir damit permanent irrational handeln. Als Antwort auf dieses Problem gibt es eine Reihe verschiedener Vorschläge, auf die wir hier nicht im Detail eingehen können.[90] Eine mögliche Lösung könnte beispielsweise darin bestehen, sekundäre Kosten zu postulieren, die der Vertrauensnehmer für den Fall trägt, dass er das Vertrauen bricht. Dies können beispielsweise materielle Sanktionen (z.B. kodifizierte Strafen bei Vertragsbruch) als auch „weiche" Sanktionen, wie etwa moralische Vorwürfe sein. Unabhängig von der begrifflich grundlegenderen Frage, ob damit wirklich genuines Vertrauen und nicht vielmehr ein bloßes Sich-Verlassen auf die Angst vor Sanktionen etabliert wird, kann dem Vorschlag entgegengehalten werden, dass in dieser Perspektive Vertrauen nicht besonders stabil erscheint. Sobald es für eine Akteurin die Möglichkeit gibt eine Sanktion zu umgehen, wird sie diese nutzen. Sie gleicht damit der von Hume eingeführten Figur des *sensible knave*[91] *,* einer Person, die sich nur solange gemäß den etablierten Regeln und Normen verhält, wie sie dadurch einen persönlichen Vorteil hat und damit der ständigen Versuchung zum Trittbrettfahren unterliegt. Darüber hinaus stellt sich auch die genealogische Frage danach, wie ein solches Sanktionssystem überhaupt unter der Bedingung radikalen Misstrauens eingeführt werden könne. Hier befinden wir uns in der Nähe des Problems, wie es überhaupt möglich sein kann aus der Hobbeschen Falle eines Naturzustandes auszutreten, ohne denselben bereits überwunden zu haben.[92]

Eine vielversprechendere Antwort auf das Problem mag darin liegen, die Annahme fallen zu lassen, dass der Interaktionspartner nur an seinem Eigennutz interessiert ist.[93] Wir haben einen Ausweg aus dem Dilemma vor Augen, wenn wir schlichtweg annehmen, dass der Vertrauenspartner unmittelbar einen Wert darin sehen kann vertrauenswürdig zu handeln (bzw. im Falle des Zeugenschaftsvertrauens die Wahrheit zu sagen). Der Vertrauensnehmer hätte dann auch unabhängig von der eigenen Kosten-Nutzen-Rechnung einen wertgebundenen Grund dafür, das ihm entgegengebrachte Vertrauen ernst zu nehmen. Dies ist im Kern die William'sche Antwort auf das Vertrauensdilemma.

Als Einschränkung betont Williams allerdings auch, dass es sich hier nur um einen ersten Schritt hin zu einer Lösung des Vertrauensdilemmas handeln kann. Damit eine Kooperation im

90 Siehe etwa die Diskussionen in: Hollis 1998; Lahno 2002, insbes. Kap.1-4; Faulkner 2017.
91 Siehe: Hume 1983, 81f.
92 Vgl. zum Konzept der Hobbeschen Falle: Baliga/Sjöström 2010.
93 Vgl. Hollis 1998, 159.

Dilemmakontext zustande kommen kann, ist nämlich auch notwendig, dass der Akteur, der vor der Entscheidung steht zu vertrauen oder nicht zu vertrauen, auch *weiß*, dass sein Interaktionspartner vertrauenswürdig ist. Nur dann würde er sich überhaupt erst dazu entscheiden zu vertrauen und es einen Anlass für eine Kooperation geben. Es muss also letztlich auch eine epistemische Bedingung[94] seitens des Vertrauenden erfüllt sein.

Nichtsdestotrotz zeigt Williams Idee zumindest eine *notwendige Bedingung* auf, die gegeben sein muss, um aus dem Vertrauensdilemma herauskommen zu können. Diese besteht, wie gesagt, in der Annahme, dass der Vertrauensnehmer jenseits seiner Überlegung zur Steigerung des egoistischen Nutzens unmittelbar einen Wert darin sieht vertrauenswürdig zu handeln. Und in ebendiesem Zusammenhang nun spricht Williams von einem „intrinsischen Wert" oder einem „inneren Wert" von Vertrauenswürdigkeit. So gelte aus Sicht des Vertrauenswürdigen: „trustworthy behavior […] has an *intrinsic value*, that it is a good thing (many things being equal) to act as a trustworthy person acts, just because that is the kind of actions it is."[95]

Diese Redeweise kann aber deshalb Anlass zur Verwirrung sein, da Williams gerade nicht zu sagen beabsichtigt, dass Vertrauenswürdigkeit einen *Selbstzweck* darstellt.[96] Ganz im Gegenteil versucht Williams ja den Wert von Vertrauenswürdigkeit dadurch zu plausibilisieren, dass er dessen *funktionale* oder *instrumentelle* Bedeutung zur Etablierung einer kooperativen Praxis des Teilens von Wissen herausstellt. Der Wert von Vertrauenswürdigkeit ergibt sich daraus, dass durch allgemeines vertrauenswürdiges Handeln der Naturzustand überwunden werden könne. Entsprechend gelte auch im Hinblick auf Vertrauenswürdigkeit (bzw. Wahrhaftigkeit): „[It] is necessary (or nearly necessary) for basic human purposes".[97]

Hier scheint sich nun aber ein Widerspruch zu ergeben. Zum einen nämlich liefert Williams eine funktionale oder instrumentelle Begründung des Wertes von Vertrauenswürdigkeit. Sie ist letztendlich deshalb wertvoll, da sie kooperative Beziehungen ermöglicht. Auf der anderen Seite jedoch qualifiziert er den Wert von Vertrauenswürdigkeit als wesentlich intrinsisch und nicht instrumentell. Dieser Widerspruch ergibt sich m.E. aus einer verwirrenden Äquivokation bezüglich des Wortes „instrumentell". Williams scheint dieses nämlich im Sinne von „am Eigennutz interessiert" zu verstehen. Er will betonen, dass sich genuine Vertrauenswürdigkeit bzw. Wahrhaftigkeit nicht als Orientierung an *eigenen Interessen* zeigen kann.[98] Damit würde uns

94 Siehe auch Williams (1988) mit weiteren Überlegungen zu den Bedingungen gelingender Kooperation.
95 Williams 2002, 90.
96 Siehe zu diesem Punkt auch: Hartmann 2011, 190; Owens 2017, 215.
97 Williams 2002, 92.
98 Vgl. insbesondere: ebd., 95.

schließlich eine zufriedenstellende Option fehlen, um zu erklären, wie wir im tatsächlichen Leben aus dem Vertrauensdilemma aussteigen können. In einer anderen Bedeutung des Begriffes „instrumentell“ verweist dieser jedoch darauf, dass Kooperation nicht *selbstzweckhaft* ist. Dessen Wert lässt sich vielmehr durch seinen Nutzen plausibilisieren. Wenn diese zwei Verwendungsweisen des Wortes „instrumentell“ auseinandergehalten werden, löst sich auch der vermeintliche Widerspruch in Williams Konzeption auf.

Warum genau Williams so emphatisch auf das Attribut „intrinsisch“ verweist, wo er meines Erachtens auch schlichtweg von einem Wert *simpliciter* sprechen könnte, hängt womöglich damit zusammen, dass das zentrale dilemmatische Problem letztlich ein Problem der tatsächlich handlungswirksamen *Motivation* des Einzelnen darstellt. Deswegen betont Williams auch, dass es nicht reicht bloß kognitiv einzusehen, dass intakte Vertrauensverhältnisse einen Wert *haben*. Der Vertrauensnehmer muss zudem diesen Wert akzeptieren und in selbstverständlicher Weise für sein eigenes Handeln leitend machen. Wir brauchen dabei keine höherstufigen Gründe dafür vertrauenswürdig zu handeln: „[to be trustworthy] makes sense to us from the inside“.[99] Die Rede vom Intrinsischen könnte deshalb auch den Zweck haben diesem Punkt einen gewissen Nachdruck zu geben. Diese soll zum Ausdruck bringen, dass der Wert von Vertrauenswürdigkeit motivational wirksam vom Vertrauensnehmer internalisiert werden muss.[100] Damit ist aber gerade nicht ausgeschlossen, dass man sich aus einer reflexiven Perspektive diesen Wert deutlich machen kann, indem man seinen funktionalen Stellenwert skizziert.

Neben dieser potenziell missverständlichen Verwendungsweise des Titels „intrinsischer Wert“ in Bezug auf Vertrauen und Vertrauenswürdigkeit bei Williams, wird auch in einem anderen Sinne vom inneren Wert von Vertrauen gesprochen. Dieser ergibt sich aus der Beobachtung, dass sowohl, wenn einem jemand vertraut, als auch, wenn sich jemand einem gegenüber

99 Williams 2002, 92.

100 Vereinzelt suggeriert Williams allerdings auch eine gewisse Unabhängigkeit von funktionalen Betrachtungen: „[...]other things being equal, it is better to be honest rather than not“ (ebd., 60). In solchen Formulierungen scheint sich eher der Gedanke auszudrücken, dass man auch vollkommen *ungeachtet der Handlungsfolgen* wahrhaftig sein sollte, Wahrhaftigkeit also ungeachtet der Konsequenzen ein Gut darstellt. Stellen wir uns etwa einen fiktiven Fall vor, in denen es keinerlei Unterschied in den Konsequenzen macht, ob ich jemanden belüge oder nicht, und auch ausgeschlossen ist, dass die Wahrheit jemals ans Tageslicht kommt. Ein solcher Fall hätte eine Parallele zu denen, die David Owens als Fälle von „*bare wrong*“ bezeichnet (Owens 2011). Die These vom intrinsischen Wert kann demnach nun auch so verstanden werden, dass es auch in diesem Fall besser sei die Wahrheit zu sagen.
An den entscheidenden Stellen jedoch führt Williams die Idee eines intrinsischen Wertes, wie gesagt, wesentlich in Abgrenzung zur Orientierung am Eigennutz ein (etwa: Williams 2002, 95). Diese Idee dient ursprünglich dazu das spieltheroetische Dilemma zu lösen. Zu sagen, dass der Wert ungeachtet der Konsequenzen besteht, würde aber noch weiter gehen als zu sagen, dass er die Eigeninteressen der Akteure übersteigt. Hier offenbart sich m.E. also im Argumentationsgang von Williams eine Erklärungslücke. Es ist nicht ersichtlich, wie sich die These begründet, dass Vertrauenswürdigkeit ungeachtet aller Konsequenzen ein Gut darstellt, denn diese folgt nicht logisch aus seinem Lösungsvorschlag für das spieltheoretische Dilemma.

vertrauenswürdig verhält, dies jenseits des funktionalen Nutzens als *in sich* wertvoll erfahren werden kann.

In diesem Zusammenhang scheint es mir auch begrifflich eingängiger als bei Williams von einem intrinsischen Wert zu sprechen. Dabei ist mit dem Begriff hier nicht gemeint, wie er auch zuweilen verstanden wird, dass etwas nicht-relational, also unabhängig von unseren Einstellungen[101] diesem gegenüber, wertvoll ist. Vielmehr geht es hier um einen Begriff des Intrinsischen, der sich wesentlich vom Begriff des Instrumentellen als funktional Begründetem abgrenzt. Oben hatten wir definiert, dass insofern etwas instrumentell wertvoll ist, dieses allein als Mittel oder Bedingung zum Erreichen eines externen Zweckes valorisiert wird. Ein Werkzeug beispielsweise ist insofern wertvoll, als es zum Reparieren taugt. Es scheint jedoch schnell deutlich zu werden, dass nicht alle Werte in diesem Sinne instrumentelle Werte sein können, da man ansonsten in einen infiniten Begründungsregress käme. Manches muss also in sich selbst, und nicht bloß als Mittel zum Zweck, wertvoll sein.[102]

Im Übrigen ist hier zu beachten, dass es sich bei dieser Unterscheidung zwischen „intrinsisch“ und „instrumentell“ um keine bloß sortale Unterscheidung handelt, sondern sie besser als *aspektuale* aufgefasst wird. So kann man sich Güter vorstellen, die *sowohl* intrinsisch wertvoll sind, *als auch* einen instrumentellen Wert haben. Wenn man Vertrauen also einen intrinsischen Wert zuschreibt, schließt dies noch nicht aus, dass Vertrauen *neben* diesem *auch* Mittel für andere Zwecke sein kann. Tatsächlich verhält es sich wohl so, dass Vertrauen und Vertrauenswürdigkeit niemals einfach nur intrinsischen Wert haben. Es handelt sich dabei nie um *bloße* Selbstzwecke. Vielmehr geht es im Vertrauen immer auch um etwas anderes, ein bestimmtes Vertrauensgut, dass der Vertrauende im Blick hat – etwa, dass ihm der andere bei der Heuernte hilft. Wir vertrauen niemals einfach so, da Vertrauen an sich eine gute Sache sei.[103]

Dennoch kann festgestellt werden, dass Vertrauen und Vertrauenswürdigkeit zuweilen nicht nur als instrumentell nützlich gesehen, sondern *auch* als in sich wertvoll erfahren werden können. Jenseits der Vorteile, die Vertrauen in Form seiner Nützlichkeit mit sich bringt, kann nämlich Vertrauen als in sich wertvolle Form der *Anerkennung* oder *Wertschätzung* empfunden werden.[104] Das gleiche gilt auch dafür, wenn unser Vertrauen ernst genommen und vertrauenswürdig

101 Siehe Lemos (2005, 3-6) für eine ausführliche Darstellung der sogenannten traditionellen Theorie intrinsischer Werte.

102 Auf diesem theoretischen Wege lässt sich die *Idee* eines intrinsisch Wertvollen als eines Regressstoppers einführen. Siehe diese Argumentation bereits bei Aristoteles: NE, 1094a21. Vgl. auch: Rønnow-Rassmussen/Zimmerman 2005, xv-xvii.

103 Vgl. Hartmann 2011, 190.

104 Siehe etwa: O‘Neill 2012, 311; McLeod 2020, Abschnitt 3.

behandelt wird. Dies ist offenkundig insbesondere im Kontext intimer Beziehungen wechselseitigen Vertrauens, wie etwa Freundschaften der Fall. In diesem Zusammenhang spricht beispielsweise Richard Holton davon, dass Vertrauen als in sich wertvoll angesehen werden kann: „I trust you; in so doing our relationship moves a little further forward. This can itself be something I value.“[105]

Dies soll nicht heißen, dass Vertrauen und Vertrauenswürdigkeit hier ein bloßes Mittel sind, um gute persönliche Beziehungen zu führen, sprich, dass beide letztlich doch wieder nur in einem instrumentellen Sinne werthaft sind. So vertraut man einer Person nicht deshalb *weil* man mit dieser eine freundschaftliche Beziehung führen will. Vertrauen und Vertrauenswürdigkeit können hier nicht in einer Zweck-Mittel-Beziehung zu einer Beziehungsform gedacht werden. Es ist vielmehr so, dass sie einen *konstitutiven* Bestandteil einer persönlichen Beziehung darstellen.

Dabei ist darauf hinzuweisen, dass intrinsisch nicht mit bedingungslos zu verwechseln ist. Davon auszugehen, dass Vertrauen etwa in intimen Beziehung pauschal wertvoll sei, käme vielmehr einer naiven romantisierten Vorstellung gleich. Auch dieser Wert ist vielmehr an bestimmte Bedingungen geknüpft, die auch kritisch geprüft werden können. Dazu gehört etwa, dass kein exploitatives Ziel mit dem Vertrauen verknüpft wird. So hat Vertrauen etwa keinen intrinsischen Wert mehr, wenn es innerhalb einer Beziehung zu manipulativen Zwecken eingesetzt wird.[106] [107]

Mit der Frage nach dem Stellenwert von Vertrauen und Vertrauenswürdigkeit im Zusammenhang mit persönlichen Beziehungsformen wie Freundschaften, werden wir uns unten noch ausgiebig beschäftigen. Wir können es deshalb an dieser Stelle bei den kurzen Erläuterungen zum intrinsischen Wert in diesem Zusammenhang belassen. Das grundlegende Ziel dieses Kapitels war es einen Überblick über die axiologischen Zusammenhänge bezüglich Vertrauen und Vertrauenswürdigkeit zu geben. Nebenbei haben wir auch eine sprachliche Unklarheit im Zusammenhang mit der Redeweise von einem „intrinsischen Wert“ im Vertrauen aufheben können. Unabhängig von der genauen *Art* der Begründung für den Wert im Vertrauen, ist für den weiteren Argumentationsverlauf allerdings im Folgenden erst einmal entscheidend, dass wir offenbar annehmen können, *dass* es plausibel ist von einem Wert in Vertrauen und Vertrauenswürdigkeit zu sprechen. Von dieser These ausgehend können wir uns wieder unseren beiden leitenden Fragen

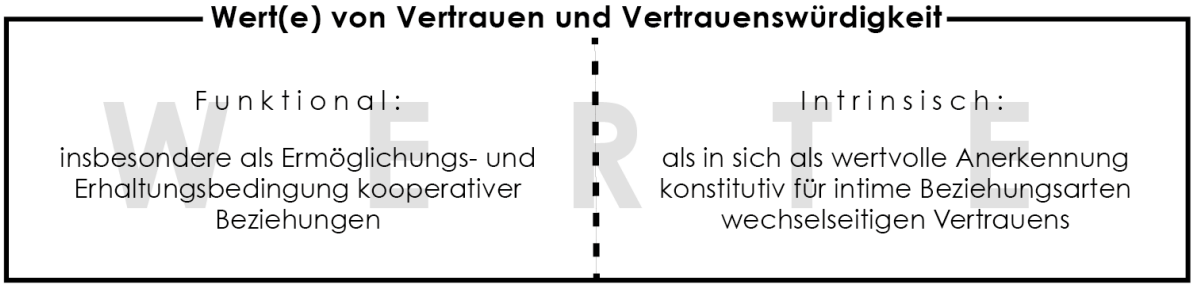
105 Holton 1994, 68.

106In der BBC-Serie *Die Schlange* ist dies ein zentrales Motiv. Der Protagonist, Serienmörder Charles Sobhraj, erschleicht sich dabei systematisch das Vertrauen seiner Opfer, indem er ihnen gegenüber selber tiefes Vertrauen, auch in symbolisch aufgeladenen Akten, zur Schau stellt.

107So kann Vertrauen etwa auch sozialen Druck erzeugen und exploitativ eingesetzt werden. Ausführlicher hierzu: Dormandy 2020.

zuwenden, bezüglich der normativen Erwartung im Vertrauen und den Gründen richtiger Art für Vertrauenswürdigkeit.

Die Ergebnisse der Analyse der Rede von einem Wert im Vertrauen und in Vertrauenswürdigkeit sei zuletzt noch einmal in einer übersichtlichen Grafik zusammengefasst:



Graphik 2: Die Werte von Vertrauen und Vertrauenswürdigkeit

4.2 Normativität, Handlungsgründe und der Wert von Vertrauenswürdigkeit

Wir werden nun auf die beiden Leitfragen dieses Kapitels zurückkommen: Wie genau lässt sich die normative Erwartung des Vertrauenden erklären und Vertrauenswürdigkeit charakterisieren? Auf beide Fragen lässt sich eine Antwort formulieren, die von einer These ausgeht, die wir aus dem vorangegangen Unterkapitel aufgreifen, namentlich der These, dass funktionierende Kooperationsbeziehungen ein Gut darstellen, für das vertrauenswürdiges Verhalten eine wesentliche Gelingensbedingung darstellen. Diese Annahme verspricht erstens eine geeignete Begründung der normativen Erwartung im Vertrauen bereitzustellen. Folgende Hypothese drängt sich hier auf: Der Vertrauenspartner sollte deshalb das Vertrauen nicht brechen, da funktionierende Kooperationsbeziehungen ein bedeutendes Gut darstellt. Vertrauen zu enttäuschen oder gar zu missbrauchen zerstört dieses Gut, Vertrauen zu erfüllen fördert dieses. Der Vertrauende erhebt die normative Erwartung gegenüber dem Vertrauenspartner demnach im Namen dieses Gutes.

Die These vom Wert der Kooperationsbeziehungen liefert zugleich eine mögliche Antwort auf die Frage, wie genau eigentlich eine vertrauenswürdige Einstellungen zu charakterisieren ist. Hier legt sich die zweite Hypothese nahe, dass eine genuin vertrauenswürdig eingestellte Person den Wert kooperativer Beziehungen unmittelbar anerkennt und ihrem praktischen Verhalten leitend macht. Dies ist letztlich auch die These, die Bernard Williams nahelegt.

Insgesamt verweisen also die Fragen, wie die normative Erwartung im Vertrauen zu

erklären ist und was eine vertrauenswürdige Einstellungen ausmacht, beide auf einen Wert von Kooperationsbeziehungen. Es scheint plausibel zu sein, einen Zusammenhang zwischen diesem Wert, der normativen Erwartung und der vertrauenswürdigen Einstellung anzunehmen. Eine andere Frage ist jedoch, wie genau dieser Zusammenhang zu spezifizieren ist. Dies ist notwendig, denn, wie wir gleich sehen werden, sind auf diesem Gebiet Fehlauffassungen möglich. Eine derselben werden wir in Form der von Thomas Scanlon sogenannten teleologischen Werttheorie diskutieren. Diese fasst den Zusammenhang zwischen Wert und normativer Erwartung bzw. der vertrauenswürdigen Einstellung gewissermaßen *zu geradlinig* auf. Dadurch verzerrt sie die Natur der normativen Relation der beiden Vertrauenspartner und das Wesen von Vertrauenswürdigkeit.

Bevor wir auf diese Diskussion genauer eingehen, beachte man, dass wir mit der zweiten Hypothese gewissermaßen unter der Hand auch auf ein bestimmtes Paradigma praktischer Rationalität referieren. Wir beziehen uns dabei auf die Idee, dass sich der Vertrauensnehmer in seinem Handeln von einem Wert leiten lässt. Diese Vorstellung weicht von der oben beschriebenen Idee praktischer Rationalität ab, nach der Handeln wesentlich der Realisierung von Präferenzen dient. Die Idee folgt also einem wertbezogenen als einem präferenzbezogenen Paradigma praktischer Vernunft, das ganz unabhängig von der Vertrauensthematik seine Vorzüge hat. Darauf gehen wir im folgenden Exkurs genauer ein.

4.3 Exkurs: Von der humeanischen zur wertbezogenen Rationalitätskonzeption

Ganz unabhängig von der Vertrauensthematik hat eine wertbezogene Theorie praktischer Rationalität seine Vorzüge gegenüber einer präferenzbezogenen. So ist das präferenzbezogene Rationalitätsparadigma in ein strukturelles Dilemma verstrickt, das mit der Frage zusammenhängt, wie genau die Rolle von Präferenzen eigentlich zu verstehen ist. Diese können entweder in einem *starken* Sinne, als *vorgegebene* Variablen eines rationalen Kalküls, verstanden werden; so haben wir schließlich oben auch die humeanischen Theorien rekonstruiert. Diese Sichtweise führt jedoch zu einer begrenzten Konzeption praktischer Rationalität und gleichermaßen zu einem unplausiblen Bild menschlichen Handelns. Als Reaktion auf diese Kritik können Verfechter des humeanischen Ansatzes den Stellenwert von Präferenzen in ihrem Ansatz abschwächen, und diese nicht mehr als praktisch-rationaler Überlegungen vorgegeben, sondern selber solchen Überlegungen *unterstehend*

auffassen.[108] Damit jedoch entleert sich das humeanische Paradigma und ist auf eine unabhängige Theorie praktischer Rationalität angewiesen.

Gehen wir zur genaueren Erläuterung zunächst auf den ersten Ast dieses Dilemmas ein. Wie wir im obigen Exkurs gesehen haben, gehen humeanische Theorien praktischer Rationalität typischerweise von einem vorgegebenen Bündel an akteursbezogenen Wünschen (*desires*) aus, die eine *fundamentale* Stellung in praktischen Überlegungen einnehmen. Gemäß dem damit skizzierten Bild rationalen Handelns fungieren Wünsche als vorgegebene Determinanten, aus denen nach Maßgabe bloß instrumenteller Überlegungen, vernünftige Handlungen folgen.[109] Dieses Bild ist jedoch deshalb unplausibel, da sich die Rationalität menschlichen Handelns auch daran bemisst, an *welchen* Präferenzen sich dieses in einem gegebenen Kontext orientiert. Man denke beispielsweise an Anscombes Beispiel einer Person, die alle grünen Bücher im Haus aufsucht und auf dem Dach verstreut.[110] Wenn man sie fragt, warum sie dies getan habe und die Antwort erhält „Ich wollte es einfach" würden wir dies kaum als zufriedenstellende Antwort akzeptieren. Vielmehr würden wir vielleicht so etwas entgegnen wie: „Offensichtlich wolltest du das. Meine Frage lautet aber eigentlich: Warum *sollte* man so etwas wollen?"

Präferenzbezogene Ansätze sind darin limitiert, dass sie keine Standards zur Einordnung der Vernünftigkeit von Präferenzen selber ausweisen können. In ihrer Perspektive ist Handeln rational, solange es instrumentell rational vollzogen wird, unabhängig davon welche Präferenzen verfolgt werden. Dann aber muss sie etwa auch die Verfolgung pathologischer Präferenzen einer Person als vernünftiges Handeln ausweisen.

Damit kommen wir nun zum zweiten Ast des Dilemmas der humeanischen Theorien. Verfechter der humeanischen Theorie können konzedieren, dass man zum Bewerten der Rationalität von Präferenzen letztlich auf präferenzunabhängige Maßstäbe angewiesen ist. Sie mögen jedoch darauf hinweisen wollen, dass letztendlich alle praktischen Überlegungen an einem konativen Moment anknüpfen müssen. Letztendlich muss die vernünftige Handlung vom Akteur schließlich auch in einer basalen Form gewollt sein, sonst würde er nämlich gar nicht handeln. Es scheint, dass die humeanische Theorie also doch etwas Wesentliches über praktische Rationalität und menschliches Handeln festhält. Mit diesem Zug jedoch ist festzustellen, dass sich die humeanische Theorie zu einer einzigen analytischen Aussage verengt, nämlich derjenigen, dass jede Form von Handlung ein konatives Moment enthält. Dies ist trivial, denn Handlungen stellen

108 Raz (vgl. 1999, 109 f.) zur strukturgleichen Unterscheidung zwischen *schwachem* und *starkem Wollen.*
109 Vgl. McDowell 1995, 78.
110 Anscombe 1957, 26f.

einen intentionalen Akt dar, also müssen sie im einem basalen Sinne *gewollt* sein. Damit liefert der humeanische Ansatz keine *substanzielle* Erläuterung mehr zur Struktur und den Standards praktischer Überlegungen.

Eine solche substanzielle Erläuterung liefern hingegen wertbasierte Theorien praktischer Rationalität. Deren Grundidee kann ausgehend von der Feststellung motiviert werden, dass praktisches Überlegen wesentlich darauf abzielt eine Antwort auf die Frage zu finden, *was man tun sollte*. Die Antwort auf diese Frage muss damit aber in irgendeiner Hinsicht die *Güte* einer Handlung ausweisen. Die *Sollens*-Frage wird nicht allein dadurch beantwortet, dass man schlichtweg auf bestehende psychologische Zustände des Akteurs verweist, sondern dadurch, dass plausibilisiert wird, dass die Handlung es *wert* ist ausgeführt zu werden. In der (neo-)aristotelischen Perspektive wird Handeln deshalb zuweilen als *konstitutiv* unter einer Hinsicht des Guten (*sub specie boni*) stehend betrachtet.[111] Das heißt, als Handelnder versteht man sich als unter dem Anspruch stehend, dass die eigene Handlung gut ist. Wann immer jemand etwas nicht einfach nur aus Versehen, sondern absichtsvoll macht, macht er dies, da er etwas Gutes in der Handlung sieht. Die Handlung ist damit an irgendeinem Wert oder Gut orientiert: Ich grüße die Nachbarn aus Freundlichkeit, ich nehme das Frühstück zu mir zum Zwecke des eigenen Wohlergehens, ich teile die Aussage p, die ich für wahr halte mit, da ich in der Wahrheit einen Wert sehe. Der Anspruch ist dabei, ein *subjekttranszendierender*[112]*:* Er besteht darin, *dass* die Handlung gut *ist*, und mir nicht nur so erscheint. Gemessen an den tatsächlichen Verhältnissen kann er entsprechend mehr oder weniger gut erfüllt sein.

Praktische Überlegungen und praktische Gründe wiederum zielen darauf ab, diesen Anspruch einzulösen. Sie sind in diesem Zusammenhang folgendermaßen einzuordnen[113]: Sie zeigen ganz grundlegend auf, *inwiefern* eine Handlung in einem konkreten Kontext tatsächlich gut ist. Praktische Gründe *sprechen für* die Handlung *im Sinne eines bestimmten Gutes*. Sie sind, wie Elizabeth Anscombe es ausdrückt, „Erwünschbarkeitscharakterisierungen“[114], das heißt, sie charakterisieren die Handlung als eine wertvolle Handlung (die insofern erwünschbar ist). Werde ich gefragt, warum ich die Hand hebe, kann ich den Grund nennen „Weil ich jemanden grüßen wollte“. Dieser Grund charakterisiert meine Handlung – bringt sie unter eine bestimmte Beschreibung, wie man mit Anscombe auch sagen könnte – , sodass ersichtlich wird, inwiefern sie

111Vgl. die ausführliche Darstellung in: Kertscher 2014.

112 Vgl. McDowell 1995.

113Die wertbezogene Theorie praktischer Rationalität bzw. praktischer Gründe basiert üblicherweise auf einer Rekonstruktion der aristotelischen Handlungstheorie. Neben Anscombe (1957) werden ihr prominenterweise auch Joseph Raz (1999) und Thomas Scanlon (1998) zugeschrieben.

114 Anscombe 1957, 73.

einem bestimmten Gut dient – hier etwa dem freundlichen Miteinander. Dabei wird in der alltäglichen Praxis des Gebens und Nehmens von Gründen meistens nicht explizit ausgesprochen welchem Gut genau eine bestimmte Handlung förderlich sein soll. Dies ergibt sich meist implizit aus dem gegebenen Kontext und wird von den Gesprächspartnern in der Regel nur expliziert, wenn es tatsächlich Missverständnisse diesbezüglich gibt.

Natürlich geben wir in *einigen* Situationen auch Präferenzen als Handlungsgründe an, etwa wenn man gefragt wird, warum man sich an der Eisdiele für Erdbeereis statt Vanilleeis entschieden hat. Hier sieht es so aus, als ob die eigene Präferenz allein ausschlaggebend sei, um die Rationalität der Handlung zu verdeutlichen: „Erdbeereis ist mein Lieblingseis". Tatsächlich aber unterstellt man in einer solchen Antwort bereits stillschweigend die Hintergrundannahme, dass es (in diesem Beispiel) besser sei, diejenige Option zu wählen, die der eigenen Präferenz entspricht. Hier orientiert man sich also wieder an einem Wert, und zwar dem Wert des eigenen Wohlergehens. Die Präferenz ist damit nur ein zu berücksichtigender *Faktor*, aber kein *Orientierungsmaßstab* praktischer Rationalität.

Insgesamt wird im Zuge dieser Kritik auch die ontologische Trennung zwischen Präferenzen und Überzeugungen fragwürdig. Zumindest wird es problematisch weiterhin die huemanische Trennungs- und Trägheitsthese aufrechtzuerhalten, nach der zwischen *bloß* konativ zu verstehenden Elementen und *bloß* kognitiv zu verstehenden Elementen zu trennen sei, wobei erstere qua ihrer ontologischen Beschaffenheit durch keine praktische Überlegung beeinflussbar seien. Wenn wir jedoch den Sinn praktischer Überlegungen wesentlich als Beantwortung der Frage verstehen, was wir tun sollten, müssen wir unsere Handlungsmotive, als durch an Gütern orientierten praktischen Überlegungen abwandelbar verstehen.

4.4 Die teleologische Werttheorie: Eine instruktive Gegenposition

Die entscheidende These, mit der wir uns im Folgenden auseinandersetzen werden, bezeichnen wir mit Thomas Scanlon als *teleologische Werttheorie.*[115] An ihr lässt sich instruktiv aufzeigen, wie der Zusammenhang zwischen dem Wert funktionierender Kooperationen, der normativen Erwartung an den Vertrauensnehmer und von Vertrauenswürdigkeit gerade *nicht* gedacht werden sollte. So vermag es die teleologische Werttheorie nicht bestimmten phänomenalen Besonderheiten der

115 Wir könnten sie auch Moore'sche Werttheorie nennen (vgl. Wallace 2005, 26).

normativen Beziehung zwischen Vertrauensgeber und -nehmer gerecht zu werden.

Die teleologische Werttheorie lässt sich in zwei Schritten darstellen. In einem ersten Schritt hält die Theorie fest, dass Werte sich wesentlich in bestimmten Sachverhalten oder Zuständen der Welt manifestieren. In den Worten Scanlons gelte hier: „The primary bearers of value are states of affairs or, over time, ways the world might go“[116]. Diese Idee an sich wäre aber nicht hinreichend, um das praktisch Relevante von Werten deutlich zu machen. Sie stehen schließlich in einem Zusammenhang zu unseren Handlungsgründen, den die teleologische Werttheorie in einem zweiten Schritt angibt. Der Vorschlag hat dabei einen klassisch konsequentialistischen[117] Einschlag, indem er die Folgen einer Handlung in den Vordergrund rückt: Demnach ist eine Handlung unter der Hinsicht eines bestimmten Wertes insofern gut, als sie darauf abzielt, (möglichst viele) Manifestationen dieses Wertes in der Welt hervorzubringen. Somit weisen praktische Gründe eine Handlung als eine solche aus, die in der Konsequenz möglichst gute Weltzustände[118] mit sich bringt. Und möglichst gute Weltzustände sind in diesem Falle solche, die *maximal werteverkörpernd* sind.

Die dahinterstehende *zustandsbeszogene Maximierungslogik*[119] hat dabei eine gewisse Plausibilität auf ihrer Seite. Für bestimmte Typen von Werten und in bestimmten Situationen scheint diese zutreffende Beschreibungen unserer ethischen Intuitionen zu liefern. Wenn ich beispielsweise durch ein wenig ärztliche Hilfe die Schmerzen einer Person lindern kann, scheint offensichtlich: Der Wert meiner Handlung hängt wesentlich damit zusammen, dass der Zustand gelinderter Schmerzen evaluativ höher liegt als der vorherige Zustand, in dem Schmerzen bestanden. Hier scheint eine offenbar triviale Grundintuition jeder konsequentialistischen Sichtweise greifbar, die Philippa Foot folgendermaßen auf den Punkt bringt; „that it can never be right to prefer a worse state of affairs to a better.“[120]

Insgesamt weist dieser Ansatz zwei bemerkenswerte strukturelle Eigenschaften auf. Zum einen gilt, dass es sich bei der teleologischen Werttheorie, um eine akteursrelative[121] Theorie

116 Scanlon 1998, 79

117Unter konsequentialistischen Ansätze können hier solche Ansätze verstanden werden, die sich, um die Frage anzugehen, was eine Akteurin in einer Situation normativ zu tun hat, allein auf eine standortunabhängige Bewertung der Sachverhalte bezieht, die eine Handlung hervorruft (Vgl. Wallace 2010, 521, Fn 19).

118„What we have reason to do, on this view (at least as far as questions of value are concerned), is to act so as to realize those states of affairs that are best—that is, have the greatest value“ (Scanlon 1998, 80).

119Samuel Scheffler spricht hier allgemein von einer „maximizing conception of rationality“ (Scheffler 1985, 414).

120Foot 1985, 199.

121Die begriffliche Unterscheidung zwischen „akteursneutral“ und „akteursrelativ“ stammt ursprünglich aus der Auseinandersetzung mit dem Utilitarismus. Sie wird meines Wissens erstmals eingeführt in: Scheffler 1985. Akteursrelative Gründe enthalten dabei „an ineliminable reference to the agent for whom they are reasons (like «that it will keep a promise I made», «that it will avoid harm to others [i.e., people other than me]» and so on). Agent-neutral reasons can be stated without such a reference: «that it would prevent some pain from occurring to

handelt. Demnach wird eine Handlungsabsicht *allein* daran gemessen, auf welche allgemeinen Weltzustände überhaupt diese abzielt. Zweitens weist der teleologische Ansatz eine grundlegend additive oder aggregierende Logik auf, die sie im Übrigen mit der Rational Choice-Theorie teilt. Während im Fall der Rational Choice-Theorie eine Handlung insofern als rational zählt als diese in ihren absehbaren Konsequenzen die Präferenzen der Akteurin *maximal* befriedigt, bestimmt sich in der teleologischen Perspektive die Rationalität einer Handlung dadurch, dass diese darauf abzielt maximal viele wertverkörpernde Weltzustände hervorzubringen.

Zieht man dieses Bild nun jedoch heran, um erstens die Natur der normativen Erwartung des Vertrauensgebers gegenüber dem Vertrauensnehmer zu beleuchten, sowie zu erklären, was Vertrauenswürdigkeit im Wesen ausmacht, dann stößt man schnell auf Ungereimtheiten, wie wir im Folgenden zeigen werden.

4.5 Zwei Einwände gegen die teleologische Lesart

Es wird schnell deutlich, dass das eben gezeichnete Bild der tatsächlichen normativen Situation zwischen den beiden Vertrauenspartnern nicht gerecht wird. Ein erstes grundlegendes Defizit der teleologischen Perspektive liegt darin, dass diese es nicht vermag den *normativen Anspruch oder die normative Erwartung* als einen *relationalen* Anspruch der Vertrauenden *gegenüber* dem Vertrauenspartner einzufangen. Genau genommen hat die teleologische Werttheorie bereits ein Problem mit dem Charakter normativer Anforderungen überhaupt: Sie qualifiziert lediglich eine Handlung als gut, und damit als evaluativ ausgezeichnet. Auf dieser Grundlage aber kann sie nicht einmal einen Unterschied zwischen supererogatorischen Handlungen – Handlungen, die man nicht machen muss, die aber lobenswert sein können – und normativ *geforderten* Handlungen – Handlungen, die man tun sollte, oder gar muss – begründen.[122] *A fortiori* steckt in der teleologischen Werttheorie nicht das artikulatorische Potenzial um auszudrücken, dass die Vertrauende *gegenüber* dem Vertrauenspartner einen normativen Anspruch hat.[123]

Um genauer nachzuweisen, dass ein solches relationales Erwartungsverhältnis besteht,

someone (or some being)»“ (Darwall 2006, 6).

122 Ebenso meint auch Philip Nickel (2007), dass das Erfüllen von Vertrauen keine rein supererogatorische Handlung darstellt. Sein Umkehrschluss jedoch, dass Vertrauen mit der Unterstellung einer (moralischen) *Pflicht* des anderen einhergeht, scheint mir zu stark (siehe dazu genauer Kapitel 5.2).

123 Ich lasse hier die metaethische Frage offen, ob solche Anforderungen *wesentlich und ausnahmslos* in dieser Form relational sind (dafür argumentierend: Wallace 2019).

mache man sich etwa Folgendes klar. In der akteursneutralen Perspektive der teleologischen Werttheorie, in der sich die evaluative Bewertung einer Handlung allein an den intendierten Konsequenzen bemisst, müsste sich eine alltagssprachlich typische Reaktion im Falle verletzten Vertrauens – wie etwa: „Du hast mein Vertrauen missbraucht/mich im Stich gelassen/etc." – übersetzen lassen können in eine unpersönliche Aussage wie etwa: „Du hast Vertrauen/die Möglichkeit gelingender Kooperation/etc zerstört". Dies klingt jedoch nicht nach einer natürlichen Reaktion auf einen Vertrauensbruch und provoziert damit eine erste noch genauer zu entschlüsselnde Anfangsirritation. Eine ähnliche Irritation lässt sich mit Bezug auf Vertrauenswürdigkeit aufzeigen. Nehmen wir an, wir fragten eine Person B, warum sich diese eigentlich gegenüber Person A so vertrauenswürdig verhält. Es klänge nun wiederum unpassend wenn B antwortete: „Weil funktionierende Kooperationsbeziehungen für mich einen wichtigen Wert darstellen und ich ihn gerne befördern möchte". [124]

In einer ersten leitenden Vermutung scheint die Ursache dieser Irritation darin zu liegen, dass die Beziehung zur anderen Person auf eine merkwürdige Weise instrumentalisiert und als bloßer möglicher *Träger* eines Wertes verstanden wird. Die Beziehung der beiden Personen erscheint in der teleologischen Perspektive als ein bloß opportuner Umstand, um kooperative Beziehungen als einen gewissermaßen unpersönlichen Wert zu fördern. Gleichermaßen scheint der Vorwurf „Du hast Vertrauen/Möglichkeiten gelingender Kooperation zerstört" irritierenderweise die konkrete vertrauende Person als einen Umstand vorzustellen, *in Bezug auf den* die beschuldigte Person einen Fehltritt begangen hat. Nennen wir dies das *Problem der Akteursrelativität* der teleologischen Werttheorie.

Diese Überlegung zeigt, dass das Tadelnswerte an Vertrauensbrüchen nicht in erster Linie darin liegt, dass damit werthafte Weltzustände nicht vermehrt wurden. Vielmehr kommt zum Vorschein, dass der vertrauenden Person im Vertrauensbruch wesentlich etwas *angetan* wurde. Dies soll nicht heißen, dass dem Vertrauenden dadurch eine bloß materielle Verletzung, sprich der risikogebundene Schaden, zugefügt wurde. Mehr noch wurde ihr wesentlich eine *normative* Verletzung zugefügt, das heißt, der Vertrauenspartner ist einer berechtigten Erwartung der vertrauenden Person *ihm gegenüber* nicht gerecht geworden. Es ist dieser Modus des Jemanden-etwas-Unrechtes-Antun oder Jemandes-Erwartung-Verletzen, der sich nicht in der Perspektive der

124 Wir müssen uns damit allerdings nicht darauf festlegen, dass derartige an einen Wert appellierende Reaktionen in einem rigorosen Sinne *falsch* sind. Vielmehr klingen sie zunächst *unangemessen* und *irritierend*. Deshalb sollten wir nicht gänzlich abstreiten, dass mit der Idee der Wertschätzung von Vertrauen *auch* die Idee der allgemeinen Maximierung von Vertrauensbeziehungen verbunden ist. Die Kritik besteht vielmehr an einem Alleingeltungsanspruch der teleologischen Werttheorie, die, wie in den obigen Beispielen durchscheint, Irritationen erzeugt ohne für diese eine Erklärung liefern zu können.

teleologischen Werttheorie repräsentieren lässt. Gemäß dieser Überlegung ist also die vertrauende Person nicht einfach nur ein Umstand in Bezug auf den es gilt richtig zu handeln, sondern sie ist eine normative Quelle von Ansprüchen oder Erwartungen.

Vielleicht mag man hier einwenden, dass auch dritte, unbeteiligte Personen den gleichen Vorwurf wie die Vertrauende aus einer anderen Perspektive erheben könnten – etwa in der Formulierung „Du hast *sie* im Stich gelassen". Dies ist tatsächlich möglich und damit scheint es, als hätte die Vertrauensperson gar keine so hervorgehobene Position für reaktive Einstellungen, sondern wäre nur eine unter mehreren möglichen Personen, die der vertrauten Person normativ defizientes Handeln vorwerfen kann. Wir werden auf diesen Punkt im nächsten Kapitel zurückkommen. Für den Moment gilt es aber zunächst nur deutlich zu machen, dass der Vertrauensperson dennoch eine gegenüber dritten Personen *hervorgehobene* Position für reaktive Einstellungen zukommt.[125] Dies liegt nicht nur intuitiv nahe, da sie immerhin diejenige ist, die vom Verhalten des Vertrauensnehmers letztlich auch persönlich betroffen ist. Es zeigt sich auch handfester an dem Umstand, dass es bestimmte alternative reaktive Einstellungen gibt, zu denen tatsächlich allein die Vertrauensgeberin berechtigt zu sein scheint. Eine solche bestünde beispielsweise darin, auf Vorwürfe gegenüber dem Vertrauensnehmer zu verzichten und ihm stattdessen seinen Vertrauensbruch zu verzeihen. Eine andere Form bestünde beispielsweise darin, Wiedergutmachungen des Vertrauensnehmers zu akzeptieren. Im Übrigen zeigt sich die hervorgehobene Stellung der Vertrauensgeberin auch für den Fall, dass das Vertrauen erfüllt wurde, insofern sie etwa Reaktionen wie Dankbarkeit gegenüber dem Vertrauensnehmer zeigen kann.

Darüber hinaus ist auch die Maximierungslogik der teleologischen Werttheorie angreifbar. Wie wir gesehen hatten, folgte die teleologische Werttheorie der Idee, dass werthafte Weltzustände als Resultat der Handlung maximiert werden sollen. Stellen wir uns, um dagegen zu argumentieren, jedoch folgende Konstellation vor: Angenommen ein Akteur hätte unter bestimmten Umständen die Möglichkeit dafür zu sorgen, dass andere Vertrauensnehmer ihr Vertrauen einhalten können oder auf andere Weise mehrere funktionierende Kooperationsbeziehungen in der Welt zu stiften indem er das Vertrauen seines Vertrauenspartners enttäuscht. Gemäß der teleologischen Werttheorie ließe sich die Güte seiner Handlung anhand der aufgerechneten Konsequenzen bestimmen und wäre womöglich *in summa* als gut anzusehen. Tatsächlich können wir uns in einem solchen Szenario aber nicht nur vorstellen, dass es fraglich ist,

125 Der Vorwurf „Du hast mein Vertrauen verletzt" ist nicht einfach nur ein Fall des akteursrelativen Vorwurfes: „Du hast jemandes Vertrauen verletzt", nur das *zufälligerweise* derjenige, der den Vorwurf erhebt, und die Person, deren Vertrauen verletzt wurde, die gleiche Person wären.

ob sich der Vertrauensnehmer mit seinem Verhalten als besonders vertrauenswürdig erwiesen hat. Vielmehr können wir uns auch vorstellen, dass der spezifische Vertrauenspartner, dessen Vertrauen enttäuscht wurde, Vorwürfe an den Vertrauensnehmer adressiert. Er scheint sich dabei (zumindest *prima facie*[126]) nicht davon beirren lassen zu müssen, dass die Handlung nach teleologischem Kalkül eine insgesamt gute Handlung war. Immerhin wurde schlichtweg *sein* Vertrauen enttäuscht, was ihn berechtigt gegenüber seinem Vertrauenspartner Vorwürfe zu erheben. In der Möglichkeit solcher Vorwürfe aber drückt sich offenbar aus, dass der Vertrauensnehmer in bestimmter Hinsicht normativ defizient gehandelt hat. Dieser Hinweis legt also insgesamt eher folgendes Bild nahe: Die Tatsache, dass der Vertrauensnehmer in einer bestimmten Vertrauensbeziehung *situiert* ist, in deren Rahmen die Vertrauensgeberin einen Anspruch hat, gibt dieser Beziehung auch einen hervorgehobenen Status im Vergleich zu anderen bloß möglichen Vertrauens- und Kooperationsbeziehungen *in der Welt überhaupt*. Die konkrete Vertrauensbeziehung lässt sich in den praktischen Überlegungen des Vertrauensgebers nicht kommensurabel mit jeder möglichen Vertrauensbeziehung als Resultat der Handlung aufrechnen. Vielmehr sollte ihr in seinen praktischen Überlegungen besonderes Gewicht zukommen. Wir können dies das *Problem des Rigorismus* der teleologischen Werttheorie nennen.

Insgesamt zeigen diese Überlegungen, dass die Normativität im Zusammenhang interpersonellen Vertrauens von der Relation der beiden Vertrauenspartner her gedacht werden muss. Der konkreten Beziehung zwischen beiden kommt besonderer normativer Stellenwert zu. Dies meint hier wiederum zweierlei: zum einen hat der Vertrauensnehmer eine spezifische Verantwortung *gegenüber* dem Vertrauensgeber, weshalb wiederum diese eine ausgezeichnete Position bezüglich möglicher reaktiver Einstellungen hat. Zum anderen sollte das *Situiertsein* des Vertrauensnehmers in einer Vertrauensbeziehung mit einer spezifischen Person, dieser eine besonderes Gewicht im Vergleich zu bestehenden oder möglichen Vertrauensbeziehungen *in der Welt überhaupt* verleihen.

Zuletzt wollen wir an dieser Stelle noch einen kurzen Blick zurückwerfen. Mit unserer Argumentation haben wir bisher eine erste Anfangsplausibilität dafür gewonnen, dass das teleologische Schema der Werttheorie aufzugeben ist, und stattdessen die Relation der beiden Vertrauenspartner als geeigneterer theoretischer Ausgangspunkt angesetzt werden muss, um die normative Situation zwischen den beiden Vertrauenspartnern zu begreifen. Mit der Kritik an der

126Es ließe sich diskutieren, ob dies Umstände sind, die das Verhalten des Vertrauensnehmers verständlich und damit entschuldbar machen. Es bliebe dann aber dabei, dass der Umstand, dass sein Verhalten *gegenüber* dem Vertrauensgeber entschuldigt werden muss, anzeigt, dass er ihr gegenüber eine *spezifische* Verantwortung hatte.

teleologischen Werttheorie ist damit aber noch nicht zwangsläufig die Ansicht widerlegt, dass Vertrauen bzw. Vertrauenswürdigkeit einen Wert hat. Die Beziehung dieses Wertes in Bezug auf die normative Erwartung des Vertrauensgebers bzw. die praktischen Gründe des Vertrauensnehmers ist in der teleologischen Perspektive jedoch falsch konzeptionalisiert.

Damit ist soweit also die Relation der beiden Vertrauenspartner als vielversprechender Ausgangspunkt für die weitere Untersuchung markiert. Um genauere systematische Einsichten in den Charakter normativen Situation zwischen den beiden Vertrauenspartnern zu gewinnen, werden wir uns im Folgenden mit dem Konzept der sogenannten relationalen Normativität auseinandersetzen und überprüfen, inwiefern sich die darin vorgenommenen theoretischen Anstrengungen zur Beschreibung der normativen Dimension zwischen den Vertrauenspartnern nutzen lassen.

5. Die relationale Normativität zwischen den Vertrauenspartnern

Im vorangegangenen Kapitel konnten wir zeigen, dass sich in Form der teleologischen Werttheorie eine zunächst naheliegende, aber letztlich problematische Vorstellung der normativen Erwartung an den Vertrauensnehmer, sowie von Vertrauenswürdigkeit nahelegt. Als entscheidenden Kritikpunkt an dieser konnten wir herausarbeiten, dass sie von der besonderen normativen Relation der beiden Vertrauenspartner abstrahiert. Diese genauer zu beschreiben wird die Aufgabe des nun folgenden Kapitels sein. Auf dieser Grundlage kann auch ein plausibleres Verständnis von Vertrauenswürdigkeit gewonnen werden.

Das entscheidende Problem der teleologischen Sichtweise bestand darin, die vertrauende Person als einen bloßen *Umstand* aufzufassen, den man sich zunutze machen sollte, um den Wert des Vertrauens zu realisieren. Demgegenüber haben wir jedoch auf den besonderen Stellenwert der vertrauenden Person gegenüber dem Vertrauenspartner aufmerksam gemacht. Sie hat einen Anspruch dem Vertrauensnehmer *gegenüber*, der ihr das Recht zu bestimmten reaktiven Einstellungen gibt. Damit scheint die vertrauende Person eine gewissermaßen *direktere* Relevanz für die normative Dimension der Beziehung der beiden zu haben.

Im Folgenden werden wir genauer und systematischer erläutern, wie die normative Struktur zwischen den beiden Vertrauenspartnerinnen zu verstehen ist. Dabei stützen wir uns maßgeblich auf einen Aufsatz von Michael Thompson, der die Relationalität moralischer Pflichten beschreibt

(5.1). Insbesondere werden wir dabei auf das Verhältnis von normativen Prinzipien und der relationalen Perspektive auf dieselben eingehen. Von diesem Punkt aus lassen sich zwei weitere Aspekte in Bezug auf Vertrauen gut beleuchten: Erstens der normative Kontext von Vertrauen und zweitens der Begriff der Vertrauenswürdigkeit. Vertrauen hat seinen Platz, wie wir zeigen werden, nur innerhalb eines vorgängigen normativen Kontextes (5.2). Vor dem Hintergrund der Theorie werden wir anschließend eine Klärung des Begriffes der Vertrauenswürdigkeit vornehmen (5.3).

5.1 Relationale Normativität

Um die spezifische Form der normativen Beziehung zwischen den Vertrauenspartnern genauer einordnen und erläutern zu können, beziehen wir uns auf den Ansatz Michael Thompsons[127] zu sogenannten bipolaren Pflichten (*bipolar obligations*) und daran anknüpfenden Überlegungen. Dabei sind bipolare Pflichten durch ein wesentlich relationales Verpflichtungsverhältnis charakterisiert, in dem eine bestimmte Person ein Recht *gegenüber* einer anderen bestimmten Person hat bzw. letztere ersterer *gegenüber* verpflichtet ist.

Am Ausgangspunkt der Überlegungen Thompsons steht die Diagnose, dass die etablierte Moralphilosophie blind für diesen relationalen Aspekt sei und moralische Zusammenhänge primär unter dem Aspekt ihrer sogenannten *monadischen Normativität* betrachte. In letzterer Perspektive wird das Handeln einer Person allein auf moralische Prinzipien bezogen und an diesen in seiner moralischen Qualität bewertet.[128] Menschliches Handeln ist entsprechend in dem Maße gut oder gerecht, wie es einem solchen moralischen Prinzip gerecht zu werden vermag, und es ist schlecht oder ungerecht, sofern es dies nicht vermag. In dieser Betrachtungsweise gerate Thompson zufolge jedoch aus dem Blick, dass sich moralische Zusammenhänge auch durch einen *relationalen* Aspekt in Form eines speziellen Nexus zwischen *Personen* auszeichnen. Dieses Phänomen markiert er nun mit dem Begriff der *bipolaren* oder auch *dikaiologischen* Normativität. Der Unterschied zwischen

127Thompson 2004. Siehe auch: Gilbert 2004. Ich beziehe mich hier primär nicht auf den Begriff der Zweitpersonalität, wie er maßgeblich von Stephen Darwall (2006) geprägt wurde. Dies hängt damit zusammen, dass Darwall mit dem Begriff der Zweitpersonalität ein weitreichenderes metaethisches Projekt verbindet, dass darin besteht Moralität entgegen drittpersonalen (etwa utilitaristischen) und erstpersonalen, (neo-)kantianischen Ansätzen als fundamental zweitpersonal auszuweisen. Auf diese Debatte will ich mich in dieser Arbeit aber nicht einlassen. Gleichwohl greife ich vereinzelt auf Ideen von Darwall zurück, wo sie helfen die normative Struktur von Vertrauen und Vertrauenswürdigkeit zu explizieren.

128Thompson betont, dass dieser Blick auf Moral nach Aristoteles und Thomas von Aquin lediglich den moralphilosophischen Bereich abdecke, der in der Tradition Aristoteles‘ bzw. Thomas von Aquins als *to nominon* oder *lex* bezeichnet werden kann: „that is, ‘what is lawful’ or ‘law’. (Here, the idea of law is, I think, to be taken very broadly, as covering inter alia any principles of what we would call morality.)“ (Thompson 2004, 338).

monadischer und bipolarer Normativität kann etwa am Fall des Versprechengebens illustriert werden, der bei Thompson und der an ihn anschließenden Diskussion durchaus als Paradefall fungiert. So steht ein Versprechensgeber A mit der Abgabe eines Versprechens in der monadischen Betrachtungsweise allein unter der Pflicht das Versprechen einzuhalten: Seine Handlung wird an dem Prinzip gemessen, dass man Versprechen im Allgemeinen nicht brechen soll. Zugleich besteht aber auch ein besonderer normativer Zusammenhang zwischen den beiden involvierten Personen, dem Versprechensgeber A und der Person B, gegenüber der A das Versprechen gegeben hat.

Diese besondere bipolare Relation zwischen den beiden Versprechenspartnern besteht darin, dass der Versprechensgeber eine Pflicht speziell gegenüber dem Versprechensnehmer hat oder, was gleichbedeutend ist, dass der Versprechensnehmer ein Recht speziell gegenüber dem Versprechensgeber besitzt. Dieser Zusammenhang drückt sich zugleich auch in einem Umstand aus, der auch titelgebend für den Aufsatz Thompsons ist: Wenn A sein Versprechen brechen würde, täte er nicht nur etwas Unrechtes *simpliciter.* Vielmehr würde er auch dem Versprechensnehmer B *ein Unrecht antun,* oder in Thompsons Worten: *he would wrong him*.[129]

Das Thema der dikailogischen Normativität klingt dabei verstreut in verschiedenen historischen Vorgängerversionen an, die Thompson systematisch und formalisiert zusammenführt. Das Subthema der Pflichten *gegenüber* anderen Personen identifiziert Thompson etwa in Überlegungen von Aristoteles und Thomas von Aquin unter den Titeln *to dikaion* bzw. *ius*. Diesem steht die Behandlung der monadischen Normativität in Form des *to nomimon* bzw. *lex* gegenüber. Ein anderer Bezugspunkt Thompsons in Bezug das Thema der Rechte gegenüber anderen Personen stellen die rechtsphilosophischen Überlegungen Wesley Hohfelds dar.[130]

Es ist entscheidend zu sehen, dass im Rahmen eines bipolaren Nexus die involvierten Personen einen ganz bestimmten normativ relevanten Status füreinander haben. Insbesondere stellt aus der Perspektive der verpflichteten Person die andere Person nicht einfach nur einen zu berücksichtigten Umstand dar. Vielmehr kommt dieser die Position zu, die zuweilen auch als *normative Autorität*[131] gegenüber der verpflichteten Person bezeichnet wird. Diesen Punkt

129Die drei einzelnen Aussagen („A hat eine Pflicht gegenüber B“, „B hat ein Recht gegenüber A“, „A würde B ein Unrecht antun“) stehen nach Thompson in einem komplexen begrifflichen Schema zusammen. Das heißt, sie drücken alle verschiedene Aspekte des letztendlich gleichen Sachverhaltes aus, nämlich, dass A und B in einem bipolaren Nexus zueinander stehen. Stephen Darwall spricht in ähnlicher Weise von einem Zirkel holistisch miteinander zusammenhängender Begriffe (*claim, authority, second-personal reason, accountability*), die sich nur unter wechselseitiger Bezugnahme erklären lassen (Darwall 2006, 11-15).

130Zudem spricht Thompson die Kritik von Thomas Scanlon und A.I. Melden an Rawls Theorie der Praxis des Versprechens an (Thompson 2004, 346). Demnach betrachte Rawls lediglich den monadischen Aspekt von Versprechen und kann dadurch nicht dessen bipolare Dimension erklären.

131Siehe die Verwendung dieses Begriffes in: Darwall 2006.

illustriert Ariel Zylbermann[132] in einer beispielbezogenen Gegenüberstellung besonders klar: Wenn ich mich etwa entgegen meinem Versprechen dir gegenüber nicht um deine Zimmerpflanzen kümmere, dann mache ich zwar *in Bezug* auf die Zimmerpflanze etwas Falsches, aber ich tue der Zimmerpflanze damit *kein Unrecht* an. *Dir* hingegen, als einem nicht bloß zu beachtenden *Faktor,* sondern einem Träger von relevanten Ansprüchen, *tue ich ein Unrecht an*. Dies verdeutlicht die hervorzuhebende *Position* der anderen Person, als normative Autorität mir gegenüber, die sich nicht einfach als besonderer zu berücksichtigender situativer Umstand übersetzen ließe. Die Auffassung einer Person als lediglich zu berücksichtigenden Faktor stellte genau unser Problem, sowohl an der oben referierten Spieltheorie, als auch an der Werttheorie, dar. So wurde die Vertrauensperson im Rahmen der Spieltheorie allein als zu berechnender Faktor zur Maximierung von Präferenzen gesehen. Für die teleologische Werttheorie stellte sie allein das Material zur Realisierung eines Wertes dar.

Bei dieser Art von normativer Autorität handelt es sich natürlich nicht um eine auf Zwangsgewalt beruhende *de facto* Autorität. Eine normative Autorität im hier relevanten Sinne zeichnet sie sich vielmehr dadurch aus, dass sie berechtigte Ansprüche gegenüber einer anderen Person hat. Dass eine Person B gegenüber einer anderen A eine solche Autorität besitzt, manifestiert sich auch in anderen Phänomenen, auf die Stephen Darwall hinweist.[133] Daraus ergeben sich nämlich ganz spezifische Interaktionsmöglichkeiten zwischen den beiden Personen: So hat allein B etwa die Möglichkeit, einzuwilligen, A von seiner Pflicht zu befreien. Bei einer Pflichtverletzung wäre B diejenige Person, gegenüber der eine Entschuldigung adressiert werden kann, und nur B wäre in der Position, eine solche Entschuldigung zu akzeptieren. B wäre zudem möglicherweise in der besonderen Position, Wiedergutmachungen von A zu verlangen etc.[134] In all diesen möglichen Zusammenhängen drückt sich die spezifische Autorität von B gegenüber A aus.

Wenn wir hier mit Thompson eine Unterscheidung zwischen monadischer und bipolarer Normativität ziehen, ist zu beachten, dass es sich dabei in erster Linie um eine Unterscheidung von zwei *Betrachtungsweisen* handelt. Wie das Beispiel des Versprechengebens zeigt, soll mit der Unterscheidung nicht gesagt sein, dass man im konkreten Fall entweder nur unter einem monadischen Prinzip oder in einem bipolaren Zusammenhang stehen kann. Im Gegenteil ist zu beachten, dass eine dikaiologiosche Relation nach Thompson stets eine *einfache* moralische Norm voraussetzt. Der Versprechensnehmer etwa kann seine besondere Position dem

132 Zylbermann 2014, 154.
133 Darwall 2011, 260.
134 Siehe auch: Wallace 2019, 86-95.

Versprechensnehmer gegenüber nur insofern rechtfertigen, als allgemein gilt, dass *man* seine Versprechen halten muss. Diese monadische Norm *begründet* die Rechtmäßigkeit des normativen Anspruches bzw. die normative Autorität des Versprechensnehmers und damit also auch, den bipolaren Nexus überhaupt. Folgerichtig gilt umgekehrt auch, dass derjenige, der ein Versprechen bricht damit stets in zwei Hinsichten etwas Falsches tut: Er tut damit niemals nur dem Versprechensnehmer etwas an, sondern tut damit zugleich auch etwas, was man im Allgemeinen nicht tun sollte, nämlich ein Versprechen zu brechen.

Dass stets eine solche moralische Norm zugrundeliegt erklärt auch, weshalb ebenso unbeteiligte Dritte jemandem, der ein Versprechen bricht, einen Vorwurf machen können. Diese haben zwar nicht die soeben geschilderte besondere Position des Versprechensnehmers gegenüber dem Versprechensgeber. Dennoch sind sie im Namen der allgemeinen Norm, dass Versprechen nicht gebrochen werden dürfen, zu moralischen Vorwürfen gegenüber der Person, die ein Versprechen bricht, berechtigt.[135]

Es gilt dieser Thompsonschen Analyse folgend also insgesamt drei Instanzen zu unterscheiden: Einem monadischen Prinzip, den situativen Umständen, auf die sich dieses Prinzip bezieht, und der Beschreibung der Relation zweier Personen. Vor diesem Hintergrund können wir klarer sehen, worin das Problem der obigen teleologischen Theorie war. Diese war blind für den letzteren Aspekt. Die zweite Person konnte deshalb nur in Form eines bloßen Elementes des Anwendungskontextes in Erscheinung treten.

Ein für unsere weitere Diskussion wichtiger Punkt ist zudem, dass bipolare Relationen nicht zwangsläufig als *deontologische* Relationen gedacht werden müssen. Die moralische Version des bipolaren Nexus, etwa im Falle eines Versprechens, stellt zwar sicherlich das leitende Paradigma im Diskurs um relationale Normativität dar. Tatsächlich aber könne die normative Qualität eines bipolaren Nexus nach Thompson verschiedenartig ausfallen. Die *normative* Relation

135 Siehe genauer hierzu: Darwall 2012. Darwall trifft diese Unterscheidung anhand des von Peter Strawson beschriebenen Unterschiedes zwischen *persönlichen* und *stellvertretenden* (*vicarious*) reaktiven Einstellungen. Persönliche reaktive Einstellungen können dabei nur von den *beteiligten* Personen eingenommen werden, oder solchen, die sich mit diesen hinreichend identifizieren. Wiederum paradigmatisch steht hierfür der Fall, in dem eine Person A Vorwürfe gegenüber B erhebt, die ein *exklusiv ihr* gegenüber abgegebenes Versprechen gebrochen hat. Die typische reaktive Einstellung in einem solchen Fall bestünde dabei im Übrigen in einer Form von Ärger (*resentment*).
Stellvertretende reaktive Einstellungen hingegen können auch von *unbeteiligten* Dritten eingenommen werden und drücken sich typischerweise nicht in Ärger, aber in *moralischen* Vorwürfen aus. Beispielhaft sei der Fall genannt, in dem Person C gegenüber B den Vorwurf erhebt, das Versprechen gegenüber A gebrochen zu haben und damit einer moralischen Pflicht nicht nachgekommen zu sein. In solchen unpersönlichen moralischen Beziehungen trete, so Darwall, der Adressierende als ein *unbeteiligter Vertreter der moralischen Gemeinschaft* auf. Als solcher fordert er schlichtweg die überindividuelle Geltung moralischer Prinzipien ein. Insofern diese allgemein gelten und er lediglich als Fürsprecher derselben auftritt, präsupponiert er, dass der von ihm geäußerte Anspruch ebenso von jeder anderen (zweitpersonal kompetenten) Person erhoben werden könne.

zwischen den beiden Personen muss sich also nicht zwangsläufig als eine moralische, im Sinne einer deontologischen, manifestieren. Die Beziehung zwischen den beiden Interaktionspartnern muss also nicht zwangsläufig in Begriffen von Recht oder Pflicht wiedergegeben werden. So könne es sich dabei ebenso auch um einen auf *Etikette* beruhenden normativen Zusammenhang handeln, der spezifiziert, was für eine Person „schicklich" wäre einer anderen gegenüber zu tun. Ebenso mag man etwa an freundschaftskonstitutive Normen denken, die angeben, was Freundinnen wechselseitig voneinander erwarten können. In diesen Fällen handelt es sich zwar um eine normative, aber nicht um eine in dem Sinne moralische Relationen, dass sie sich in deontologischen Kategorien beschreiben lassen könnte. Freunde etwa haben als solche kein *Anspruchsrecht* darauf, dass der andere sich so verhält, wie man es von einem Freund erwartet. In diesem Sinne hält Thompson für bipolare Normativität fest: „they may be shifted into various gears, or sung in various keys, […] only one of them specifically moral".[136]

Unsere obige Kritik nun an der teleologischen Rekonstruktion der Normativität interpersonellen Vertrauens legt es nahe, die Beziehung zwischen den Vertrauenspartnern in ebendiesem bipolaren bzw. relationalen Sinne zu verstehen. Im Vertrauen hat die Vertrauende also wesentlich eine *gerichtete* normative Erwartung *an* die Vertrauenspartnerin, die wiederum speziell der Vertrauenden *gegenüber* normativ gefordert ist, das Vertrauen zu erfüllen.[137] Die Normativität von Vertrauen zeigt sich insofern wesentlich in Form dieses Nexus zwischen den Vertrauenspartnerinnen.[138] Dies vorausgesetzt können aus der allgemeinen Diskussion um bipolare bzw. relationale Normativität auch direkt einige interessante Konsequenzen für den speziellen Fall interpersonellen Vertrauens gezogen werden. Eine davon betrifft das, was ich den *normativen Kontext* von Vertrauen nenne.

136Thompson 2004, 345f.

137Damit ist im Übrigen auch aus begrifflichen Gründen ausgeschlossen, dass man jemanden in Bezug auf eine supererogatorische Handlung vertrauen kann, zumindest sofern man supererogatorische Handlungen als solche definiert, die über das hinausgehen, was normativ von einem erwartet werden kann. Man kann zwar *darauf vertrauen, dass* sich jemand im supererogatorsichen Sinne wohltätig verhält. Beispielsweise bin ich prinzipiell dazu berechtigt *darauf zu vertrauen, dass* mein Hausflurnachbar mir spontan einen Kuchen backen wird, ohne dass ich davon ausgehe eine berechtigte normative Erwartung diesbezüglich gegenüber zu haben. Sofern wir eine normative Erwartung jedoch als Merkmal genuinen interpersonellen Vertrauens akzeptieren, kann ich nicht ohne Weiteres *ihm* vertrauen, dass er dies tun wird.

138Vgl. Kaminski 2020, 240-267.

5.2 Der normative Kontext von Vertrauen

Wie wir gesehen haben, setzt jede Form von relationaler Normativität immer auch voraus, dass ein bestimmtes monadisches Prinzip einschlägig ist. So begründet sich die Verpflichtung des Versprechensgebers gegenüber dem Versprechensnehmer dadurch, dass es die generische Pflicht gibt, dass *man* (*pro tanto* zumindest) sein Versprechen halten muss. Folgerichtig gilt dies ebenso für die normative Relation zwischen den beiden Vertrauenspartnern. Die normative Erwartung des Vertrauenden muss schließlich in irgendeiner Form als *angemessen* verstanden werden können. Man kann sich gewissermaßen nicht ohne Weiteres dazu *entscheiden* einer bestimmten Person in bestimmter Hinsicht ϕ zu vertrauen, und damit allein *qua Vertrauensentscheidung,* eine entsprechende normative Anforderung dieser gegenüber kreieren. Ich kann mich etwa nicht einfach dazu entschließen, meinem Hausflurnachbarn zu vertrauen, mir morgen einen Kuchen vor die Tür zu stellen und ihn somit gewissermaßen *ex nihilo* normativ daran binden. Auch wenn ich ihm dies mitteile, gibt es damit immer noch keine Rechtfertigung dafür, ihm gegenüber eine normative Erwartung einnehmen zu können, geschweige denn ihm Vorwürfe zu machen, falls er diese enttäuscht. So gesehen ist Vertrauen also nicht einfach normativ *autark*: Es kann nicht aus sich selbst heraus bestimmte normative Anforderungen erschaffen. Vielmehr scheint umgekehrt zu gelten, dass Vertrauen stets in *vorgängig bestehenden normativen Zusammenhängen situiert ist.*[139] Vertrauen findet seinen angemessenen Platz im Kontext von bestehenden normativen Anforderungen, die situationsspezifisch überhaupt erst bestimmen, in Bezug worauf es prinzipiell *möglich* ist, einer anderen Person zu vertrauen, in welchen Hinsichten also Vertrauen überhaupt eine angemessene Kategorie darstellt.

Das irreführende Bild von der Autarkie des Vertrauens könnte etwa die Vertrauensdefinition von Paul Faulkner nahelegen. Faulkner definiert: „A trusts S to Φ (in the affective sense) if and only if (1) A depends on S Φ-ing; and (2) A expects (1) to motivate S to Φ (where A expects this in the sense that A expects it of S that S be moved by the reason to Φ given by (1)).“[140] Das Problem

139Man mag hier vielleicht einen Fall wie den folgenden als vermeintliche Ausnahme konstruieren: Nehmen wir an, wir lebten in einer Welt, in der qua Konvention Versprechen nur dann gültig seien, wenn der Versprechensnehmer dem Versprechensgeber in Bezug auf das Versprechen auch vertraut. In diesem Fall bestünde die Verpflichtung des Versprechenden nicht unabhängig vom Vertrauen des Versprechensnehmers, schlichtweg da Vertrauen eine Geltungsbedingung seiner Verpflichtung darstellt. Abgesehen davon, dass dieser Fall einen konstruierten Spezialfall darstellt, trifft er aber gar nicht den hier entscheidenden Punkt. Der Punkt ist nicht, dass Vertrauen nicht Teil der normativen *Geltungsbedingungen* sein kann, sondern vielmehr, dass es immer schon bestimmte Normen voraussetzt. In dem konstruierten Beispiel wäre dies etwa die auf Konvention beruhende Norm, dass man sein Versprechen zu halten hat (sofern einem vertraut wird). Der Punkt ist: Würde eine solche Norm nicht bestehen, könnte man einer anderen Person in Bezug auf dieser nicht vertrauen.

140Faulkner 2011, 146.

an dieser Definition ist es, dass sie suggeriert, es sei *allein* die *Abhängigkeit* von A gegenüber S bezüglich der Handlung Φ, die A als motivierenden Grund von S erwartet. Allerdings stellt bloße Abhängigkeit von jemandem *als solche* nicht zwangsläufig eine Rechtfertigung für eine normative Erwartung dieser Person gegenüber bereit. Dies geht nur, wenn diese Erwartung situativ *gerechtfertigt* ist. Wenn ich vor dem langen Wochenende meine Einkäufe nicht erledige, bin ich in einem gewissen Sinne davon abhängig, dass mein Mitbewohner Lebensmittel übrig haben wird. Aber ich kann damit noch nicht ohne Weiteres von ihm *erwarten*, Lebensmittel für mich übrig zu haben. Dafür müssten zusätzliche Bedingungen ins Spiel kommen, beispielsweise, dass er mir versprochen hat, für mich mit einzukaufen. Nur in diesem Fall nämlich, ist es durch eine einschlägige Norm gerechtfertigt, dass ich von meinem Mitbewohner erwarten kann, für mich einzukaufen. Folglich ist es auch nur in diesem Falle angemessen, dass ich ihm diesbezüglich vertraue.

Ein verklärtes Bild von Vertrauen mag es erschweren, diesen Punkt zu sehen. Wenn man davon ausgeht, dass es im Normalfall etwas Positives sei, wenn einem vertraut wird, übersieht man vielleicht, dass dies tatsächlich vielleicht nicht immer willkommen ist. Gerade da interpersonelles Vertrauen mit einer normativen Erwartung einhergeht, und somit eine normative *Bürde* für den Vertrauenspartner bedeutet, unterliegt Vertrauen stets gewissen rechtfertigenden Angemessenheitsbedingungen.

Der normative Kontext von Vertrauen bestimmt im Übrigen ebenso die Adäquatheitsbedingung für die Kategorie des Misstrauens. Misstrauen ist nämlich nicht, wie man meinen könnte, einfach *nur* die *Abwesenheit* von Vertrauen *tout court.* Kants Nachbarin etwa, die sich nur auf Kants Spaziergänge verlässt, ohne ihm diesbezüglich zu vertrauen, misstraut Kant damit noch nicht. Mehr noch, es wäre ihrerseits *kategorial unangemessen* Kant in diesem Zusammenhang zu misstrauen – ebenso wie es etwa unangemessen wäre, meinem Nachbarn bezüglich des Kuchens zu misstrauen. Misstrauen ist allein die Abwesenheit von Vertrauen unter der *Bedingung, dass eine berechtigte normative Erwartung besteht.*[141] Auch die Frage, in Bezug worauf man einer Person überhaupt misstrauen kann, wird also dadurch vorgegeben, was situativ gerechtfertigterweise von einer anderen Person erwartet werden kann.

Was diese normative Einbettung von Vertrauen betrifft, lohnt es sich deren vielfältige möglichen Ausprägungen vor Augen zu führen. Hierzu soll im Rest dieses Kapitels ein Überblick gegeben werden, auch um zu vermeiden, dass unsere Konzeption in diesem Punkt als zu eng

141Dies ist die Erweiterung der Idee von Katherine Hawley (2011), (2019), die dies nur für *commitment*-bezogenes Vertrauen so festhält.

missverstanden wird. Am Ende werden wir dabei zudem noch einmal auf die Rolle von sozialen Praktiken in diesem Kontext eingehen.

Bei den normativen Anforderungen an den Vertrauenspartner können zunächst grob zwei Arten unterschieden werden: Zum einen gibt es solche Anforderungen, die auf einem vorangegangenen öffentlichen *commitment* des Vertrauenspartners beruhen. Mit einem solchen *commitment* meine ich hier einen öffentlichen Akt, durch den sich der Vertrauenspartner freiwillig an den Vertrauenden normativ bindet, beispielsweise dadurch, dass er ein Versprechen abgibt. Genauer werden dies aber unten erläutern. Auf diese Form *commitment*-bezogenen Vertrauens gehen wir gleich genauer ein. Zuvor wenden wir uns aber dem Vertrauen zu, das auch *ohne* ein öffentliches *commitment* der anderen Person bestehen kann. Zuweilen vertrauen wir nämlich auch Personen im Lichte normativer Anforderungen, die unabhängig von einem öffentlichen *commitment* der anderen Person gelten: Wenn mein Nachbar etwa ein Paket für mich annimmt, kann ich ihm vertrauen, dass er dieses nicht einfach öffnen wird, insofern ich davon ausgehe, dass man dies im Allgemeinen nicht tun *sollte* – zumindest solange es dafür keine gute Rechtfertigung gibt. Diese allgemeine Norm ermöglicht es, dass interpersonelles Vertrauen hier als überhaupt angemessene Kategorie ins Spiel gebracht werden kann. Ebenso kann ich, wenn ich mich in einer Notsituation befinde, anderen vertrauen, dass sie mir helfen werden, sofern es als normativ geboten angesehen werden kann, dass Personen in dringlicher Not zu helfen ist. Zuweilen ist, was wir von einer Person erwarten auch durch deren Rolle oder Funktionsträgerschaft mitbestimmt. Wenn wir sagen, dass wir der Briefträgerin, der Zahnärztin, oder der Politikerin vertrauen, meinen wir damit in der Regel nicht, dass wir ihnen in jeder Hinsicht vertrauen, sondern in Bezug auf die durch ihre jeweilige Rolle definierten normativen Anforderungen.

An diesen Beispielen sieht man auch exemplarisch veranschaulicht, dass für die entscheidende normative Relation zwischen den beiden Vertrauenspartnern in der Regel zunächst einmal irrelevant ist, ob man faktisch vertraut oder nicht. Mein Nachbar sollte meine Pakete nicht öffnen, *unabhängig* davon, ob ich ihm diesbezüglich vertraue, oder ob ich ihm misstraue. Vertrauen ist zwar notwendigerweise in einen normativen Rahmen eingebettet. Dessen Geltung ist aber in der Regel nicht daran gebunden, dass hier eine Person *tatsächlich* einer anderen vertraut.[142] Das schließt allerdings nicht aus, dass Vertrauen in bestimmten Kontexten von *zusätzlicher*

142Mit der Unabhängigkeit zwischen der normativen Einbettung von Vertrauen und Vertrauen selber, meine ich deshalb hier nur, dass der Umstand, dass man faktisch vertraut in der Regel keine Geltungsbedingung der situativen Normen ausmacht. Auf einer anderen Ebene mag man aber dennoch einen *begrifflichen* Zusammenhang zwischen Vertrauen und den kontextuellen Normen konstatieren. Dieser besteht eben darin, dass solche Normen den begrifflichen Rahmen spannen, in dem Vertrauen angemessen ist.

normativer Bedeutung sein kann. Wenn ich etwa das Versprechen von jemandem breche, von dem ich weiß, dass er mir diesbezüglich vertraut, habe ich nicht nur einfach meine Pflicht verletzt, das Versprechen zu halten – ich habe auch das in mich gesetzte Vertrauen verletzt. Dies scheint, zumindest unter bestimmten Bedingungen, von zusätzlicher normativer Relevanz zu sein. Um klären zu können, worin diese Relevanz besteht, müssen wir jedoch erst genauer verstanden haben, was es mit Vertrauen genau auf sich hat – wir kommen deswegen am Ende dieses zweitens Teils der Arbeit auf dieses Thema zurück. Was in jedem Fall gilt, ist, dass auch wenn Vertrauen an sich normativ gewichtig sein kann, es sich dennoch auf eine bereits bestehende und unabhängige normative Anforderung gegenüber dem Vertrauenspartner beziehen muss.[143]

Es ist wichtig zu betonen, dass, wenn wir hier von vorgängigen normativen Anforderungen sprechen, der Begriff relativ weit zu verstehen ist. So beschränkt sich der Begriff keinesfalls auf deontisch bestehende Anforderungen, schließt dieser aber gleichwohl nicht aus. Wenn es etwa darum geht, dass ich in einer dringlichen Notsituation jemandem vertraue mir zu helfen, mögen wir es mit ausgesprochen starken deontischen Ansprüchen meinerseits zu tun haben. In Modalverben ausgedrückt gilt für diesen Fall, dass die andere Person mir geradezu helfen *muss*. Auf der anderen Seite des Spektrums mag sich Vertrauen auf eher schwache normative Erwartungen beziehen, etwa solche, die zur Etikette gehören. Ich kann etwa Bekannten vertrauen, dass sie mich willkommen fühlen lassen, wenn ich sie besuche, ich kann Freunden darauf vertrauen, dass sie an meinen Geburtstag denken oder mich im Krankenhaus besuchen etc. Die Vielfalt an Färbungen der normativen Beziehungen, die Thompson in Bezug auf das Konzept der bipolaren Normativität hervorhob, ist demnach also auch im Bereich interpersonellen Vertrauens anzunehmen.

Darüber hinaus können die betreffenden Normen auch unterschiedlich sozial implementiert sein. So kann es sich etwa um soziale Normen handeln, deren Befolgung bereits eine fest anerkannte soziale Praxis darstellt. Es kann sich aber beispielsweise auch um rechtliche Normen im Kontext eines entsprechenden Institutionensystems handeln, das etwa auch mit staatlicher Autorität und einem organisierten Sanktionsapparat ausgestattet ist. Beispielsweise kann ich auch einer Person, mit der ich einen beglaubigten Vertrag abschließe, vertrauen, dass sie diesen einhalten wird. Zugegeben ist Vertrauen in solchen Kontexten typischerweise keine besonders naheliegende Kategorie. Das soll heißen: Es gibt in der Regel keinen Anlass dazu, die Beziehung etwa zwischen zwei Vertragspartnern als Vertrauensbeziehung zu charakterisieren. Dies ergibt sich zum einen daraus, dass, gerade weil wir es mit einem institutionalisierten und sanktionsgestützten

143 Vgl. O'Neill (2017), der davon ausgeht, dass Vertrauen maximal eine „second-order-obligation" stiften könne.

Kontext zu tun haben, man üblicherweise kein besonders hohes Risiko mit der Unterzeichnung des Vertrages eingeht. Ein Vertragspartner wird sich aufgrund der vorliegenden Anreizstruktur (etwa die Angst vor Sanktionen und Reputationsverlust[144]) wahrscheinlich nicht dazu entscheiden, den Vertrag zu brechen. Und selbst wenn er dies täte, wäre der erlittene Schaden dadurch begrenzt, dass einen vermutlich eine Entschädigung erwartet. Hinzu kommt, dass – wie wir unten noch genauer sehen werden – die genannten vordergründigen Anreize, den Vertrag nicht zu brechen, nicht die Motive sind, auf die sich Vertrauen in eine Person bezieht. Wir haben bereits oben verdeutlicht, dass sich Vertrauen in eine Person wesentlich auf die Vertrauenswürdigkeit der anderen Person bezieht. Vertrauenswürdigkeit zeichnet sich jedoch durch ganz bestimmte, im folgenden Unterkapitel genauer zu charakterisierende Handlungsmotive aus, die mit Sanktionsangst oder Reputationssorgen nichts zu tun haben.

Dennoch kann nicht vollkommen ausgeschlossen werden, dass auch in solchen juristischen Kontexten irreduzibel ein gewisses Risiko und die Vertrauenswürdigkeit einer Person eine Rolle spielen. Im Zusammenhang mit *realen* Kontraktverhältnissen kann auch Vertrauen und Misstrauen eine gewisse Rolle spielen: Man denke etwa allein an die Anlässe, dem anderen beim Zustandekommen des Vertrages zu vertrauen oder zu misstrauen, ebenso denke man daran, dass wir darauf vertrauen können, dass der Vertrag keine Fallstricke enthält, oder auch, dass der Vertragspartner nicht doch unter Umständen einen Vertragsbruch in Kauf nimmt etc. Vertrauen sollte also auch in juristischen Kontexten nicht aus begrifflichen Gründen als sinnvolle Kategorie ausgeschlossen werden.[145]

Typischerweise mögen wir bei Vertrauen aber gerade an Zusammenhänge denken, in denen die normative Erwartung an eine andere Person gerade nicht durch eine fest etablierte soziale Praxis oder gar ein Sanktionssystem unterfüttert ist. Dies liegt vermutlich daran, dass man in solchen Zusammenhängen unmittelbarer auf die *Vertrauenswürdigkeit* der anderen Person angewiesen ist, und damit interpersonelles Vertrauen eine einschlägigere Kategorie darzustellen scheint – wie gesagt stellt dies ein Thema dar, auf das wir unten noch genauer eingehen werden.

144Knud Eijer Løgstrup etwa merkt diesen Punkt an. Fest etablierten Normen könne man demnach aus folgenden Motiven heraus folgen: „just out of habit, or because we are afraid that the social order might break down or become unstable, or out of fear of sanctions, or in order to make ourselves look meritorious in our own eyes or those of others“ (K.E. Løgstrup zitiert nach Stern 2019, 57). Gerade wenn diese Möglichkeit jedoch *nicht* gegeben ist, spielen Vertrauen und Vertrauenswürdigkeit eine hervorgehobene Rolle.

145Damit spreche ich hier gar nicht das *genealogische* Problem an, dass Kontrakte nicht-kontraktförmiges Vertrauen entwicklungslogisch voraussetzen – ein Punkt, auf den etwa Durkheim prominenterweise hinweist (Durkheim 1988, I. Buch, Kap.7). Es geht vielmehr darum, dass es nicht begrifflich ausgeschlossen werden kann, dass man Vertragspartnern vertraut. In gewöhnlichen Kontexten handelt es sich dabei lediglich womöglich nicht um eine *naheliegende* Einstellung.

Zunächst einmal aber kommen wir noch etwas detaillierter auf diejenige Gruppe von normativen Anforderungen zu sprechen, die sich auf ein vorangegangenes öffentliches *commitment* des Vertrauenspartners beziehen. Diese Fälle rückt insbesondere Katherine Hawley[146] ins Zentrum ihrer Vertrauenskonzeption. Dabei geht Hawley, ebenso wie wir im vorangegangen Unterkapitel, von der Beobachtung aus, dass Vertrauen nicht einfach *ex nihilo* normative Anforderungen erschaffen könne. Im Besonderen richtet sie sich damit gegen solche Ansätze, die Vertrauen lediglich als eine bestimmte Form von mentalem Zustand einer Person auffassen, etwa als eine epistemische Erwartung, die mit einem bestimmten Gefühl einhergeht. Diese Sichtweise könne nicht erklären, warum Vertrauen (ebenso wie Misstrauen) *in bestimmten Situationen kategorial unangemessen* ist. Dafür müsse man vielmehr verstehen, dass Vertrauen mit einer normativen Erwartung einhergeht, die vom Vertrauenden in der jeweiligen Situation als angemessen aufgefasst wird. Als die entscheidende Angemessenheitsbedingung einer solchen Erwartung nennt Hawley nun ein vorangegangenes öffentliches *commitment* der Vertrauensperson, etwa, dass sie ein Versprechen gegeben hat. Hawley versteht dies als die eigentliche, in jedem Fall aber als eine besonders wichtige Form von Vertrauen. Unsere Diskussion von oben hat jedoch gezeigt, dass man auch im Zusammenhang mit *commitment*-bezogenen normativen Anforderungen vertrauen kann. Wir sehen hier also das *commitment*-bezogene Vertrauen als nur eine mögliche Form.

Wir werden nebenbei bemerkt den Begriff des *commitment* im Folgenden nicht übersetzen, da es für ihn keine besonders treffende Übersetzung zu geben scheint. Am nächsten kommt ihm vielleicht der Begriff der *Zusage*. Dieser weckt jedoch womöglich irreführende Assoziationen, dass wir es dabei in der Regel mit expliziten Sprechakten zu tun haben, was, wie wir gleich sehen werden, für *commitments* im hier relevanten Sinne nicht unbedingt der Fall sein muss. Außerdem verstehen wir Zusagen in der Regel auf eine zukünftige Handlung bezogen – etwa, wenn ich dir zusage, dich morgen zu besuchen. Auch dies ist für den Begriff des *commitments* hier nicht notwendig.

Wir hatten bereits erwähnt, dass diese *commitment*-basierte Form von Vertrauen nicht nur von Hawley, sondern in einem Großteil der Literatur als von besonderem Interesse angesehen wird. Dies hat meines Erachtens vor allen Dingen zwei gute Gründe[147]: Zum einen stellen die Fälle des

146Hawley 2011, 2019.

147Hawley selber nennt hierfür auch den Grund, dass das *commitment*-basierte Vertrauen passgenauer auf den Begriff der Vertrauenswürdigkeit gemünzt sei. Sie meint, Vertrauenswürdigkeit verstehen wir als die spezifische Disposition dazu seine *commitments* zu erfüllen, und nicht als Disposition dazu seinen normativen Anforderungen *überhaupt* nachzukommen. In diesem Punkt jedoch neige ich dazu ihr nicht zu folgen. Oben haben wir schon darauf aufmerksam gemacht, dass sich Vertrauen auch auf normative Anforderungen an Personen beziehen kann, die nicht auf einem *commitment* beruhen. Ebenso meine ich, dass nichts dagegen spricht die Erfüllung solcher normativen Anforderungen auch als Ausdruck von Vertrauenswürdigkeit anzusehen.

Sprechervertrauens und des Vertrauens in Versprechen immer noch die am intensivsten diskutierten Sonderformen von Vertrauen dar. Diese basieren jedoch eindeutig auf einem öffentlichen *commitment*. Zweitens, spricht für die Relevanz des *commitment*-basierten Vertrauens, dass in der Vertrauensliteratur das Phänomen des *Verrates* (*betrayal*) von Vertrauen eine vieldiskutierte Rolle spielt. Immerhin verstand bereits Annette Baiers[148] die Möglichkeit des Verrates als ein wesentliches Kennzeichen von Vertrauen. Verrat aber stellt eine ganz bestimmte normative Verletzung dar, die grundlegend im Nichteinhalten einer *Zusage* oder eines öffentlichen *commitments,* nicht in der Verletzung einer Pflicht *per se* besteht.[149] Wer mir in einer Notsituation nicht hilft, in der er unter der Pflicht stand zu helfen, der hat damit mir gegenüber zwar eine Pflicht verletzt, er hat mich aber nicht notwendigerweise verraten.

Lässt sich aber noch etwas Genaueres dazu sagen, was ein solches öffentliches *commitment* ist, und wie genau dieses Anlass für eine normative Erwartung geben kann? Im Allgemeinen kann man sagen, dass auch für Hawley das Abgeben eines Versprechens den Musterfall eines *commitments* darstellt.[150] Prinzipiell versteht sie den Begriff des *commitment* aber recht weit. Demnach handelt es sich bei einem *commitment* grundlegend um einen intentionalen und kommunikativen Akt oder eine Reihe solcher Akte, durch die eine Akteurin die normative Beziehung zu einer anderen Person bewusst umstrukturiert, genauer: durch die sie sich gegenüber der anderen Person zu etwas normativ bindet.[151] Dieser intentionale Akt kann dabei verschiedene Formen annehmen. Er muss etwa nicht notwendigerweise durch einen expliziten Sprechakt erfolgen, wie es etwa bei Versprechen und Mitteilungen der Fall ist. Je nach Kontext können auch implizite Konventionen das Verhalten einer Person als eine Art verpflichtendes *commitment* ausweisen. Wenn ich mir etwa in der Mensa eine Essensportion ausgeben lasse, willige ich damit stillschweigend ein, dieses auch (zumindest) zu bezahlen. Man kann sich zudem auch durch eine längere Reihe von Akten und eher schleichend auf bestimmte Handlungen normativ binden. Wenn ich etwa über einen längeren Zeitraum eine Freundschaft zu einer anderen Person entwickle, gehe ich damit auch gewisse Pflichten ein, die man als Freund gegenüber dem anderen Freund hat.[152]

148 Baier 1986.
149 Vgl. Hinchman 2017
150Hawely 2019, 10.
151Es stellt damit eine Form von Ausübung normativer Macht (*normative power*) dar. Siehe zu diesem Begriff in der Debatte um Versprechen: Heuer 2012; Owens 2006; Watson 2009.
152Erzählt mir der Freund etwas Privates, bin ich dazu verpflichtet dies nicht weiterzuerzählen. Nicht weil ich mich speziell *dazu* explizit bereit erklärt hätte. Vielmehr habe ich solche Arten von Pflichten auf mich genommen, indem ich die Freundschaft eingegangen bin. Hier spricht Hawley an einer Stelle auch von sogenannten meta-*commitments*: Manche *commitments*, wie das Eingehen einer Freundschaft, haben weitere potenzielle *commitments* im Gefolge (Hawley 2019, 77f.).

Weiterhin können auch aus regelmäßigen Handlungen bestimmte *commitments* erwachsen: Nehmen wir etwa an, ich verabrede mich mit einem bestimmten Freundeskreis über einen langen Zeitraum und jeden Freitagabend zum Abendessen. Auch aus diesen regelmäßigen Verhaltensweisen können stillschweigend bestimmte normative Anforderungen entstehen, beispielsweise die, dass ich den Freund:innen vorzeitig Bescheid gebe, falls absehbar ist, dass ich am Freitagabend keine Zeit haben werde.

Die Beispiele zeigen zudem, dass die normative Färbung der *commitments,* ebenso wie oben in Bezug auf nicht *commitment*-bezogenes Vertrauen, unterschiedliche Ausprägungen haben kann. Es muss sich dabei nicht zwangsläufig um deontisch oder gar juridisch ausgeprägte *commitments* handeln, etwa wie beim Versprechen oder Vertragsschluss. Das heißt unter anderem auch, dass es sich dabei nicht zwangsläufig um eine Verpflichtung handeln muss, die mit einem *Anspruchsrecht* einer Person gegenüber der anderen einhergeht. Vielmehr kann es gewissermaßen auch „weichere" Formen von *commitments* geben, wie im Fall des Eingehens einer Freundschaft, der gerade kein Anspruchs*recht* korrespondiert. Freunde können freundschaftliches Verhalten untereinander nicht einfordern oder erzwingen. Und doch bedeutet das Eingehen einer Freundschaft auch das Eingehen einer gewissen normativen Bindung.

Zu ergänzen wäre an diesem Ansatz von Hawley allerdings noch, dass *commitments* in vielen Fällen nur dann normativ bindend sind, wenn diese *auch* von demjenigen, demgegenüber das öffentliche *commitment* gemacht wird, akzeptiert werden. Ein Versprechen beispielsweise ist nur dann bindend, wenn es auch der Versprechensnehmer annimmt. Auch diese Annahme kann implizit vollzogen werden, etwa wenn ich dem Versprechensgeber zunicke oder zumindest nicht explizit sein Versprechen ablehne. Alle diese Akte wechselseitiger und freiwilliger, aber zugleich normativ bindender *commitments* können dem zugerechnet werden, was Stephen Darwall Transaktionen (*transactions*) nennt. Darunter versteht er verschiedene Formen normativer Bindungen, die allein unter der Bedingung zustande kommen, dass beide Interaktionspartner diese freiwillig annehmen: „Transactions take place at all only if there is the relevant uptake from both parties, and they presuppose a common authority both parties have to demand free participation."[153] Versprechen sind dabei nur ein Beispiel in dieser Klasse der Transaktionen. Als weitere Formen nennt Darwall etwa gemeinsame Abmachungen, die Annahme einer Einladung oder eines Gesuches etc. In allen diesen Fällen gestalten Akteure auf Basis freiwilliger Zustimmung ihre normative Situation: Sie kreieren selbstständig normative Erwartungen des einen gegenüber dem

153Darwall 2011, 269.

anderen bzw. entsprechende Verpflichtungen. Ein Großteil der für die *commitment*-Theorie von Hawley relevanten Zusagen sind in ebendiesen Kontext zu verorten.[154]

Zur Frage, was genau als ein solches *commitment* im relevanten Sinne gilt, ist zudem festzuhalten, dass dies zumindest in einem Großteil der Fälle eine Frage der sozialen Konvention ist.[155] Versprechengeben stellt eine konventionsbasierte soziale Praxis dar, zu der schlichtweg gehört, dass man einen Sprechakt wie „Hiermit verspreche ich dir ϕ zu tun" als bindende Zusage dazu ϕ zu tun versteht. Ebenso gilt die Annahme des Mensa-Essens an der Essensausgabe *vor dem Hintergrund der üblichen Praxis* als Zusage selbiges später an der Kasse zu bezahlen. Ganz grundlegend gilt, dass die beschriebenen normativen Relationen zwischen Vertrauenspartnern nicht im luftleeren Raum konstruiert, sondern im Kontext bereits bestehender, normativ relevanter sozialer Praktiken vollzogen werden. Diese stellen den praktische Orientierungsrahmen für beide Akteure dar.

Soziale Praktiken haben dabei eine grundlegende, transzendentale Rolle für *commitment*-bezogene Vertrauensbeziehungen. Allein wenn es zur Konstitution der normativen Beziehung wesentlich ein wechselseitiges *commitment* geben muss, muss zum einen derjenigen, der dieses *commitment* gibt, wissen, was er tut. Zum anderen muss derjenige, demgegenüber das *commitment* gemacht wird, diese als ein solches *commitment* identifizieren können. Dafür jedoch müssen beide Interaktionspartner schon wissen, was es *überhaupt* heißt, ein solches *commitment* zu geben. Sie müssen bereits einen Begriff eines solchen *commitments* haben, sprich sie müssen die Praxis des Gebens und Akzeptierens eines solchen *commitments* kennen. Um etwa einer anderen Person ein Versprechen geben zu können, muss man bereits wissen, was es *im Allgemeinen* bedeutet ein Versprechen zu geben. Und ebenso muss dies die Person wissen, der man das Versprechen gibt, um das Versprechen als Versprechen auffassen zu können. Beide Versprechenspartner müssen also bereits mit der *Praxis des Versprechens* vertraut sein.

Praxis kann hier somit gewissermaßen als Medium verstanden werden, das ermöglicht, dass sich zwei Personen überhaupt sinnvoll in einer gemeinsamen Tätigkeit aufeinander beziehen können. Matthias Haase spricht hier von einer „expression condition" die für das Vollziehen gemeinsamer Tätigkeiten *als gemeinsamer* Tätigkeiten erfüllt sein muss: „Since the minds related

154 Natürlich soll dies keineswegs suggerieren, dass solche Transaktionen stets im idealen Maße freiwillig vollzogen werden. Tatsächlich können sie auch „weniger freiwillig" zustande kommen, etwa wenn asymmetrische Machtbeziehungen zwischen den beiden Akteuren im Spiel sind, die sich (implizit) in deren Transaktion auswirken. Dies kann etwa in Form von Manipulation, offenem oder unterschwelligem sozialen Druck oder Ähnlichem geschehen. Unter solchen Bedingungen ist die Transaktion zumindest verdeckt exploitativ und kritikabel, und nur stark eingeschränkt freiwillig.

155 Hawley deutet diesen Punkt auch an: „that mutual expectation and convention give rise to commitment unless we take steps to disown these" (2019, 10f.).

to one another in this way are distinct individuals and thus external to each other, their meeting has to be mediated".[156] Dies gilt nicht weniger für das Eingehen von *commitments*. Übliche Praktiken erklären zudem auch, wie so viele unserer *commitments implizit* ablaufen können, ohne ausgesprochen werden zu müssen. Um erneut unser Beispiel aufzugreifen: Vor dem Hintergrund der etablierten sozialen Praxis *bedeutet* die Annahme des Mensa-Essens auch eine Verpflichtung dazu dies zu bezahlen, weshalb es hier diesbezüglich keiner ausdrücklichen Abmachungen bedarf. Das schließt natürlich nicht aus, dass im Einzelfall nicht eindeutig ist, was im Sinne der üblichen Praxis als ein *commitment* zählt und es hierbei auch zu zwischenmenschlichen Missverständnissen kommen kann. Martin Hartmann etwa betont aus diesem Grunde, dass jede Praxis des Vertrauens immer auch ein gewisses Praxisvertrauen voraussetzt: „Weil wir nie absolut sicher sein können, dass die, denen wir Vertrauen entgegenbringen, tatsächlich unserem Praxisverständnis folgen, setzt die Möglichkeit des Vertrauens so etwas wie *Praxisvertrauen* voraus."[157]

Damit wollen wir es an dieser Stelle mit der Diskussion der beiden Formen an normativen Hintergrund für Vertrauen belassen. Das grundlegende Ziel dieses Kapitels war es deutlich zu machen, dass Vertrauen (und Misstrauen) immer einen Bezug zu vorgängigen situativen normativen Anforderungen an den Vertrauenspartner aufweist. Diese spannen gewissermaßen einen kategorialen Rahmen für interpersonelles Vertrauen. Außerdem wurde ein Überblick über die Vielfalt an denkbaren normativen Zusammenhängen diesbezüglich gegeben. Damit wollen wir uns im Folgenden nun systematisch der Frage widmen, die wir im vorangegangen Kapitel bereits einmal als zweite Leitfrage ausgewiesen haben. Dies ist die Frage nach dem Wesen von Vertrauenswürdigkeit.

5.3 Vertrauenswürdigkeit: Eine tugendethische Konzeption

Über Vertrauenswürdigkeit haben wir im bisherigen Verlauf bereits einiges provisorisch angedeutet und vermutet. So sind wir etwa davon ausgegangen, dass sich diese in einer bestimmten Einstellung der Vertrauensperson zeigt, die im Zusammenhang mit der normativen Erwartung dieser gegenüber besteht. Mittlerweile sind wir in einer guten Position, uns noch einmal in systematischer und ausführlicher Form dem Thema der Vertrauenswürdigkeit zuzuwenden. Wir werden uns dabei mit den vorherrschenden Konzeptionen auseinandersetzen und für ein

156Haase 2014,123. Diese Idee, steht auch schon im Zentrum von Michael Thompsons Aufsatz (2004).
157Hartmann 2011, 25.

tugendethisches Verständnis von Vertrauenswürdigkeit argumentieren.

Dafür beziehen wir uns auch auf den gegenwärtigen Vertrauensdiskurs, in dem sich grob zwei verschiedene Typen an Konzeptionen von Vertrauenswürdigkeit identifizieren lassen. Zum einen wird die Ansicht vertreten, dass sich Vertrauenswürdigkeit im Wesentlichen in bloßen Absichten zu kooperativen Handlungen zeigt, und sich nicht durch bestimmte Motive oder Handlungsgründe einer Person auszeichnet.[158] Dies werden wir absichtsbezogene Ansätze nennen. Demgegenüber gibt es eine Bandbreite an Positionen, die bestimmte löbliche Handlungsmotive als wesentlich für Vertrauenswürdigkeit ansehen. Je nach Version dieser Ansicht werden dabei verschiedene Kandidaten an Motiven genannt, die als konstitutiv für Vertrauenswürdigkeit ausgewiesen werden. Vertrauenswürdigkeit kann sich demnach in Form moralisch integrer Motive[159], oder in Form von persönlicher Benevolenz[160] dem Vertrauenden gegenüber ausdrücken.

Die Kritik an der ersten Position haben wir oben bereits wiederholt anklingen lassen, wir werden gleich aber noch einmal ausführlicher auf diese eingehen. In Bezug auf die verschiedenen Versionen der zweiten Ansicht, stellt sich die Herausforderung angesichts der verschiedenen genannten Motive darzulegen, *was genau* diese als Motive von Vertrauenswürdigkeit charakterisiert. Es stellt sich die Frage danach, was diese verschiedenen Motive zusammenhält und es erlaubt sie unter dem Titel Vertrauenswürdigkeit zu rubrizieren.[161] Wir werden deshalb ein übergreifendes charakteristisches Merkmal von Vertrauenswürdigkeit manifestierenden Handlungsmotiven identifizieren.

Nach diesem Schritt werden wir im darauf Folgenden nachweisen, dass ein entscheidender Aspekt von Vertrauenswürdigkeit noch unbeleuchtet blieb. Von einer vertrauenswürdigen Person sprechen wir nämlich erst dann, wenn diese eine relativ robuste Disposition zum Handeln aus solchen Motiven aufweist. Vertrauenswürdigkeit zeigt sich damit insgesamt als eine gute Charakterdisposition, oder eine praktische Tugend. Zuletzt werden wir noch auf den in meinen Augen oft missverstandenen Zusammenhang zwischen Vertrauenswürdigkeit und Kompetenzen eingehen.

158 Walker 2006, 81. Darüber hinaus findet sich ein derart weites Verständnis von Vertrauenswürdigkeit auch in den üblichen spieltheoretischen Konzeptionen (etwa Hardin 2002). Wie wir sehen werden, ist auch Katherine Hawleys (2019) Verständnis von Vertrauenswürdigkeit hier einzuordnen.

159 Etwa: McLeod 2002.

160 So eine Lesart von: Baier (1986).

161 Eine ähnliche Frage stellt sich mit Bezug auf Ansätze in der empirischen Forschung, die verschiedenen Komponenten (etwa: Kompetenz + Benevolenz + Integrität) für Vertrauenswürdigkeit vorschlagen, etwa: Johnson 1999; Mayer et al. (1995). Es bleibt damit die Frage offen, wie der sie verbindende Begriff der Vertrauenswürdigkeit zu charakterisieren ist. Was hält die Elemente *zusammen*, sodass sie *als* Elemente von Vertrauenswürdigkeit anzusehen sind?

5.3.1 Absichtsbezogene Konzeptionen von Vertrauenswürdigkeit

Worin genau manifestiert sich die Vertrauenswürdigkeit einer Person? Ein Teil der Literatur geht davon aus, dass sich diese in der bloßen Absicht zu bestimmten kooperativen Handlungen zeigt – unabhängig von den genauen Motiven oder Gründen für diese. Dieser Position ist etwa auch der Ansatz von Katherine Hawley zuzurechnen. Nach Hawley manifestiert sich Vertrauenswürdigkeit in der *Absicht so zu handeln, wie es den öffentlichen commitments entspricht, die man gemacht hat*. Wohlgemerkt betont Hawley, dass es hierbei allein um die Absicht geht und nicht um die spezifischen *Motive*, aus denen heraus man diese Absicht hat. „Absicht" meint hier für sie lediglich die Absicht, *diejenige Handlung auszuführen*, die das zuvor öffentlich abgegebene *commitment* verlangt – unabhängig davon, *warum* man als Vertrauenspartner diese Absicht hat: „To be trustworthy in some specific respect, on this view, it is enough to behave *in accordance* with one's commitment, regardless of motive".[162]

Als anderer prominenter Vertreter eines solchen absichtsbezogenen Ansatzes kann Russel Hardin angesehen werden. Sein Ansatz weicht dabei in den Details und den Begriffen von Hawleys Ansatz ab. So fokussiert er dem Schwerpunkt nach nicht auf *commitments*, sondern auf kooperative Beziehungsformen. Vertrauenswürdigkeit zeigt sich entsprechend wesentlich in der Absicht kooperativ zu sein. Außerdem lässt Hardin für seinen Ansatz zwar auch mögliche Handlungsmotive variabel, betont aber, dass Eigeninteresse ein besonders starkes Motiv ist, das damit auch eine hervorzuhebende Relevanz für kooperative Beziehungen hat. Entsprechend hält er pointiert fest: „trustworthiness is just the capacity to judge one's interests as dependent on doing what one is trusted to do."[163] Im Wesentlichen versteht er Vertrauenswürdigkeit als eine Form von „enlightened self-interest"[164].

Für die verschiedenen Versionen absichtsbezogener Ansätze ergibt sich das gleiche Problem, dass sie den eigentliche Wesenskern von Vertrauenswürdigkeit verfehlen. Sofern absichtsbezogene Ansätze nämlich die dahinter stehenden Handlungsmotive offen lassen, müssen sie auch zulassen, dass bloßes Eigeninteresse ein Motiv von Vertrauenswürdigkeit sein kann – im Falle Hardins stellt dies sogar ein hervorgehobenes Motiv dar. Dies jedoch ist unplausibel, wie wir im Folgenden mithilfe dreier Argumente aufzeigen wollen:

162Hawley 2019, 76; Hervorhebung von mir. Siehe auch: „a sincere promise is one which is backed by an intention to act as promised" (ebd., 79).
163Hardin 2002, 28.
164Ebd., 66.

(1) Das Problem der Gründe falscher Art. Man stelle man sich zunächst Folgendes an unserem *Heu-Ernte-Beispiel* angelehnte Szenario vor: Angenommen Bauer A kennt Bauer B sehr gut, schätzt ihn charakterlich aber als eine durch und durch egoistisch handelnde Person ein. Unter dieser Voraussetzung würde A wohl kaum einen Vertrauensakt gegenüber B eingehen, sprich ihm in Form einer kooperativen Vorleistung bei der Ernte helfen. Nun sei aber angenommen, die Genossenschaft der Landwirte entscheidet sich dazu, jedem Bauern für nachbarschaftliche Erntehilfe einen angemessenen Lohn auszuzahlen. Vor diesem Hintergrund würde es A wohl als deutlich wahrscheinlicher einschätzen, dass B ihm im Gegenzug bei der Ernte helfen wird und deshalb vermutlich mit B kooperieren. A sieht es also als deutlich wahrscheinlicher an, dass B die *Absicht haben wird, das zu tun, was er, A, normativ von B erwarten würde*. Nach Hawleys Auffassung würde er also eine vertrauenswürdige Einstellung zeigen. Tatsächlich aber scheint es unangemessen davon zu sprechen, dass A nun B mit der Entscheidung der Genossenschaft als *vertrauenswürdiger* einschätzt. Dies zu sagen irritiert, denn an Bs *prinzipieller charakterlicher Einstellung* gegenüber A hat sich im Grunde nichts geändert. Es sind einzig und allein Bedingungen eingetreten, unter denen es wahrscheinlicher ist, dass B die Absicht haben wird, das zu tun, was von ihm erwartet wird. B (oder vielmehr Bs Verhalten) ist so gesehen *lediglich* auf eine bestimmte Weise verlässlicher, in dem Sinne, dass B erwartbar die Handlung ausführt, die von ihm gefordert ist.

(2) *Das Problem der fehlenden Löblichkeit.* Zudem[165] ist darauf aufmerksam zu machen, dass wir im Allgemeinen andere Personen für ihre Vertrauenswürdigkeit schätzen oder loben. Es handelt sich damit um eine *evaluativ positive* Disposition einer Person. Damit ist ebenfalls in Frage gestellt, dass es sich einfach um ein Handeln aus Eigennutz handeln kann. Was sollte an einer solchen Einstellung schließlich lobenswert sein?

(3) *Das Problem der Instabilität.* Der Eindruck, dass Bauer B im Grunde nicht vertrauenswürdiger geworden ist, kann auch durch die Überlegung gestützt werden, dass das kooperative Verhalten von B weiterhin strukturell instabil ist, während wir Vertrauenswürdigkeit gerade als eine im gewissen Grade stabile Disposition zu verstehen scheinen. Vom Bauern B würden wir erwarten, dass sich dieser nur solange kooperativ verhält, wie dies seinem eigenen Vorteil gereicht. Sobald er sich jedoch in einem kontingenten Kontext befindet, in dem er die

165Zudem kann man auch noch Folgendes bedenken: Es scheint plausibel anzunehmen, dass vertrauenswürdiges Verhalten kein Vertrauen enttäuschen kann. Oben aber haben wir ausgeführt, dass die Enttäuschung von Vertrauen auch mit bestimmten Motiven des Vertrauenspartners zusammenhängt (siehe Kap. 3.4.2). Wenn Vertrauenswürdigkeit jedoch vollkommen unabhängig von den Motiven definiert wird, ergibt sich hier zumindest eine Erklärungslücke. Für einen ähnlichen Punkt: siehe Hartmann 2011, 215.

Möglichkeit hat Trittbrett zu fahren, würde er zu seinem Vorteil ausscheren und das Vertrauen brechen. Die Frage, ob er kooperiert fällt damit eins zu eins mit der Frage zusammen, ob er unter den gegebenen Bedingungen eigene Vorteile aus der Kooperation ziehen kann. In dieser Hinsicht ist seine Kooperationswilligkeit also maximal kontextsensitiv und strukturell instabil. Daran würde sich auch prinzipiell nichts ändern, wenn wir seine eigennützigen Ziele als extrem langfristig ausgelegt auffassen würden, etwa wie bei der von David Hume diskutierten Figur eines scharfsinnigen Spitzbuben (*sensible knave*).[166] Insofern er die langfristigen Vorteile einer stabilen Kooperationsbeziehung im Blick hat, und nicht nur die kurzfristigen Vorteile, die trittbrettfahrende Manöver mit sich bringen, würde er sicherlich zwar öfter kooperieren. Im Prinzip bleibt der Sachverhalt aber dennoch der gleiche: Sobald er sich in einer Situation befände, in der er langfristige Vorteile durch Trittbrettfahren nicht verlöre, oder die kurzfristigen Vorteile schlichtweg eindeutig überwögen, würde er sich nicht mehr kooperativ verhalten.

Dies steht im Gegensatz zu unserem vortheoretischen Verständnis von Vertrauenswürdigkeit. Wenn wir jemandem Vertrauenswürdigkeit zuschreiben, schreiben wir ihm damit eine im gewissen Maße stabile Disposition zu: Sie ist zumindest zu einem gewissen Grad invariant gegenüber Kontexten, in denen Vertrauensbrüche eigennützige Vorteile mit sich bringen. So schiene es etwa merkwürdig inkonsistent bzw. erklärungsbedürftig, wenn ich Person P als vertrauenswürdig einschätzte, zugleich aber P nur in solchen Situationen vertraute, in denen ich sicher bin, dass sie keinen Vorteil daraus haben wird, wenn sie mein Vertrauen bricht.

Für die Gültigkeit dieser Überlegung scheint es im Übrigen keine Rolle zu spielen, von welcher Art von eigennützigen Motiven genau wir hier sprechen: Sei es das Streben nach kurzfristigem oder langfristigem materiellen Gewinn, Angst vor möglichen Sanktionen oder Reputationssorgen.[167] Diese werden zwar typischerweise allesamt als *kooperationsfördernde* Motive genannt, eben da sie die Verlässlichkeit des Verhaltens einer Person, unter bestimmten entgegenkommenden Bedingungen, steigern können. Diese machen lediglich berechenbarer, dass

166Hume 1983, 81f.

167Hinzu kommt unser obiger Punkt, dass auch *bestimmte wohlmeinende* Motive nicht kennzeichnend für Vertrauenswürdigkeit sind. Dazu gehört eben das wohlmeinende Motiv im Allgemeinen Vertrauen in der Welt zu fördern, auch wenn man damit auf rigoristische Weise den konkreten Vertrauenspartner übergeht. In ähnlicher Weise streicht Jessica Miller heraus, dass es als solches keine passende Charakterisierung einer vertrauenswürdigen Haltung darstellt, wenn man einen allgemeinen Wert in der Beziehung sieht: Nehmen wir an „a husband is motivated not to betray this trust (say, by being unfaithful) by the thought that their committed marital relationship is of value to him. But if this is true, then it is not the case that fulfilling his wife's trust is his primary concern. Rather, it is second to his concern to stay in a committed marital relationship with her. But, what is a committed marital relationship in which fulfilling trust obligations is not of primary concern?" (Miller 2000, 45).

sich die andere Person kooperativ verhalten wird. Es handelt sich bei ihnen aber nicht um Motive, die Vertrauenswürdigkeit im eigentlichen Sinne zum Ausdruck bringen.[168]

Aus dieser Überlegung ergibt sich im Übrigen auch, dass sich Vertrauenswürdigkeit nicht in der *bloßen* Pflichterfüllung einer Person zeigt, wenn wir unter Pflichterfüllung lediglich das Handeln *gemäß* vorgeschriebener Pflichten verstehen. Oben haben wir bereits darauf hingewiesen, dass Akteure, insbesondere im Kontext fest institutionalisierter und mit Sanktionsgewalt unterfütterter Normensysteme (etwa dem juridischen System), eine Reihe von Motiven haben können sich an diese zu halten. Knud Eijer Løgstrup listet folgende Motive auf, die Akteure haben können um etablierten Normen zu folgen: „just out of habit, or because we are afraid that the social order might break down or become unstable, or out of fear of sanctions, or in order to make ourselves look meritorious in our own eyes or those of others".[169] In einem bestimmten Sinne des Wortes haben Personen, die sich von diesen Motiven leiten lassen, vielleicht ihre Pflicht „erfüllt". Eine Person etwa, die ihr Versprechen aus Angst vor möglichen Sanktionen einhält, hat damit schließlich kein Recht einer anderen Person verletzt. Folgen wir unserer obigen Argumentation, dann kann man jedoch nicht sagen, dass sie sich damit schon als vertrauenswürdig erwiesen hat. Vielmehr scheint es gerade so zu sein, dass sich die wahrhaftige Vertrauenswürdigkeit einer Person gerade in den Situationen erweist, in denen sie eben keine äußeren Anreize dazu hat sich vertrauenswürdig zu verhalten.

5.3.2 Vertrauenswürdigkeit und normative Anforderungen

Die vorangegangenen Überlegungen zeigen, dass eine adäquate Theorie der Vertrauenswürdigkeit spezifizieren muss, welche Handlungsmotive diese charakterisieren. In der Literatur lassen sich bisher eine Reihe von möglichen Kandidaten ermitteln, die als typische Motive von Vertrauenswürdigkeit ins Spiel gebracht wurden. Annette Baier etwa legt die Idee nahe, dass sich Vertrauenswürdigkeit als eine Form von persönlichem Wohlwollen (*goodwill*) gegenüber dem Vertrauenden manifestiert.[170] Auch wenn der Begriff bei ihr zugegebenermaßen nicht klar definiert

168Solche Motive sind als solche zunächst nur Anreiz für Kooperationen als solchen. Eine andere Frage ist aber, ob die Stärkung solcher Motive, etwa indem man das Sanktionspotential erhöht, längerfristig auch die Bereitschaft für Vertrauen und Vertrauenswürdigkeit erhöht oder gerade unterminiert. Siehe hierzu auch die Überlegungen in: O'Neill 2002.

169K.E. Løgstrup zitiert nach: Stern 2019, 57.

170Sie legt es genau genommen nur indirekt nahe, da sie Vertrauen als ein Sich-Verlassen auf das Wohlwollen des anderen definiert (Baier 1986, 234).

wird, so schließt der Vorschlag zumindest nicht aus, dass damit so etwas wie persönliche Benevolenz gegenüber dem Vertrauenden gemeint ist. Dagegen wurde nun mehrfach eingewandt, dass dies in jedem Fall nicht das einzige Motiv von Vertrauenswürdigkeit darstellen kann.[171] Richard Holton etwa hielt Baier kritisch den folgenden Fall entgegen: „we trust our enemies not to fire at us when we lay down our arms and put out a white flag. There is not necessarily much goodwill there. Our enemies' restraint might be grudging, driven simply by a belief that it is wrong to shoot someone who has surrendered."[172] [173]

Natürlich ist es möglich den Begriff des Wohlwollens bei Baier auch in einem weiteren Sinne zu verstehen. Demnach umfasse er nicht bloß persönliche Zuneigung, sondern etwa auch die moralische Integrität, Gewissenhaftigkeit einer Person usw. Dies verspricht zwar einen Fall, wie den von Holton beschriebenen einzufangen. Die Folgefrage, die sich dann jedoch aufdrängt, lautet, was diese unterschiedlichen Einstellungen gemeinsam haben. Was an ihnen zeichnet sie als Ausdruck von Vertrauenswürdigkeit aus? Auf diese Frage hat Karen Jones eine Antwort gegeben. Auch sie sieht das Dilemma, dass der Begriff des Wohlwollens entweder zu eng verstanden werden müsse, oder als eine theoretisch konturlose „grab bag list"[174] möglicher Motive fungiere.[175] Ihr Ansatz, den sie als „simple view"[176] bezeichnet, zielt deshalb darauf ab, ein gemeinsames Merkmal all der möglichen Motive von Vertrauenswürdigkeit auszumachen, dass diese als Motive *von Vertrauenswürdigkeit* klassifizieren lässt. Diesbezüglich schlägt sie nun vor, dass Motive von Vertrauenswürdigkeit durch eine gewisse Responsivität gegenüber der Abhängigkeit des Vertrauenden bestimmt sind. Für eine genuin vertrauenswürdige Person stellt demnach der Umstand, dass die Vertrauende auf ihr zählt, *als solcher* einen Grund dafür dar, das Vertrauen zu erfüllen. Für die Vertrauenswürdige gelte: „the fact that someone is counting on you can, *all by itself* and without further incentive, activate responsiveness".[177] Hierbei hat der Modalsatz „*all by*

171 Etwa: Holton 1994, 66; Jones 2012b, 67f.; McLeod 2002, 21-23; Walker 2006, 76f..

172 Holton 1994, 66.

173 Siehe auch die Kritik daran von Carolyn McLeod (2002). McLeod weist darauf hin, dass wir insbesondere in förmlichen Kontexten beispielsweise in der Arzt-Patienten-Beziehung, vom anderen nicht unbedingt erwarten, dass uns dieser persönlich wohlgesonnen ist. Wenn Patienttinnen darauf angewiesen wärem, dass ein Arzt ihnen persönlich gegenüber wohlgesonnen ist, spricht dies insgesamt nicht für seine Vertrauenswürdigkeit. McLeod sieht deshalb in moralischer Integrität den Musterfall von Vertrauenswürdigkeit: „what we really want [people we trust] to be motivated by is moral integrity" (McLeod 2002, 23).

174 Jones 2012b, 67.

175 Ähnlich wird der Begriff der Vertrauenswürdigkeit bei Bernd Lahno (2002, 160-166) recht weit verstanden und umfasst sowohl die Bereitschaft dazu, um eines vom Vertrauenden geschätzten Gutes willens zu handeln, als auch das Handeln aus Pflicht. Was bei Lahno ausbleibt ist aber eine Erläuterung, was diese beiden Arten von Motiven strukturell gemeinsam haben, sodass sie beide Vertrauenswürdigkeit verkörpern.

176 Jones 2012b, 70.

177 Jones 2012b, 68; Hervorhebung von mir. In dieser Formulierung klingt dabei ihre Definition von Vertrauen nach: „trust is an attitude of optimism that the goodwill and competence of another will extend to cover the domain of our interaction with her, together with the expectation that the one trusted will be directly and favorably moved by the

itself and without further incentive“ die Funktion explizit Eigeninteresse als mögliches Handlungsmotiv auszuschließen. So mag es für eine bloß am Eigennutz interessierten Person zwar auch relevant sein, dass die Vertrauende auf sie zählt. Allerdings wäre dies für sie allein in der Form eines zu berücksichtigenden *Umstandes* bei der Verfolgung eigennütziger Interessen von Belange. Insofern ein Vertrauensnehmer aber *ausschließlich* an den persönlichen Vorteilen einer funktionierenden Kooperationsbeziehung interessiert ist, und sie allein im Lichte dessen die andere Person nicht im Stich lässt, handelt sie nicht vertrauenswürdig.[178]

Als entscheidenden Vorzug ihrer Charakterisierung nennt Jones, dass diese weit genug gefasst sei, um der breiten Vielfalt an Motiven gerecht zu werden, die wir unserem gewöhnlichen Verständnis nach als Motive für Vertrauenswürdigkeit akzeptieren. Der Ansatz sei neutral gegenüber der „deeper story“ weshalb jemand in der Tatsache, dass jemand auf einen zählt einen Handlungsgrund sieht. Für eine Person, die uns gegenüber vertrauenswürdig ist könne schließlich gelten: „Perhaps she harbors friendly feelings towards us; [...]. Or perhaps she is generally benevolent, or honest, or conscientious, and so on.”[179]

Auch wenn ich denke, dass der Charakterisierungsvorschlag von Jones zwar korrekt ist, so meine ich, dass er auch weiter gehen könnte. Das kennzeichnende vertrauenswürdiger Motive lässt sich meines Erachtens noch genauer spezifizieren. So scheint sich Vertrauenswürdigkeit nicht allein darin zu zeigen, dass man in der *schieren Tatsache*, dass jemand auf einen zählt, einen Handlungsgrund sieht. Was Jones aus dem Blick nimmt ist die Rolle von *normativen Anforderungen* an den Vertrauenspartner zur Bestimmung von Vertrauenswürdigkeit. Oben haben wir darauf hingewiesen, dass Vertrauen stets im Kontext von *berechtigten normativen Erwartungen* überhaupt seinen prinzipiellen Platz findet. Das gleiche scheint aber auch für Vertrauenswürdigkeit zu gelten.

Zwei kurze Beispiele können uns helfen klarer zu sehen, was der Ansatz von Jones nicht in den Blick nimmt. Der erste Fall greift dabei ein Beispiel von Hawley auf, das wir bereits erwähnten. Angenommen ich weiß, dass sich mein Bürokollege *ohne irgendeinen Anlass* dazu entscheidet darauf zu *zählen*, dass ich Champagner zur Arbeit bringe. In diesem Fall wäre ich zwar nicht responsiv im Sinne Jones, dies spräche aber wohl kaum gegen meine Vertrauenswürdigkeit. Ein ähnliches Argument lässt sich mithilfe des zweiten Beispieles entwickeln, das auf ein Szenario

thought that we are counting on her“ (Jones 1996, 4).

178Jones wendet sich damit insbesondere gegen Russel Hardins Konzeption von Vertrauensbeziehungen als Interessenbeziehungen, aus der sich eine Idee von Vertrauenswürdigkeit als bloßem „enlightened self-interest“ ergibt (Jones 2012b, 66f.)

179Jones 1996, 7. Zitiert aus: Jones 2012b, 67.

unmoralischen Vertrauens verweist. Angenommen ein Bankräuber zählt auf seinen Komplizen, der mit dem Fluchtwagen auf ihn wartet. Falls sich letzterer loyal verhält und seinen Kollegen nicht im Stich lässt, verhielt er sich zwar responsiv im Sinne Jones. Es ist jedoch höchst fragwürdig, ob wir sein Verhalten als Ausdruck seiner Vertrauenswürdigkeit auffassen würden. In jedem Fall kann man sagen, dass sein Verhalten kaum den Inbegriff von Vertrauenswürdigkeit verkörpert. Nehmen wir umgekehrt an, der Fluchtwagenfahrer würde Gewissensbisse bekommen, sich der Polizei stellen und somit die Erwartung seines Komplizen enttäuschen. In diesem Verhalten scheint eher ein Anzeichen für seine eigentliche Vertrauenswürdigkeit zum Vorschein zu kommen.

Beide Fälle weisen darauf hin, dass Vertrauenswürdigkeit nicht hinreichend durch seinen responsiven Charakter bestimmt ist. Unser Charakterisierungsvorschlag lautet deshalb vielmehr: Es ist für die vertrauenswürdige Person nicht der Umstand ausschlaggebend, dass jemand auf sie zählt *tout court*. Vielmehr misst sich ihre Vertrauenswürdigkeit daran, dass sie sich in ihrem Handeln intrinsisch an dem orientiert, was sie tun *sollte*, was in einer gegebene Situation also eigentlich normativ Geboten ist, und deshalb von ihr normativ zu erwarten ist.[180] Eine vertrauenswürdige Person handelt also aus Einsicht in die Geltung des normativ Gebotenen. Auch Vertrauenswürdigkeit ist damit kategorial stets in einem Raum normativer Anforderungen situiert.

Was genau eine Vertrauensperson in einer gegeben Situation tun sollte, was also von ihr normativ gefordert ist, kann natürlich umstritten sein. Dass hier Unterschiede in den tatsächlichen Bewertungen vorliegen können, widerspricht nicht dem soeben gemachten Vorschlag. Es bleibt dabei: Auch wenn wir in unserer Bewertung der Vertrauenswürdigkeit des Fluchtwagenfahrers vermutlich von der seines Komplizen abweichen würden; wir würden genau *deshalb* von ihm abweichen, da wir das loyale Verhalten desselben im Widerspruch zu dem sehen, was er eigentlich, moralisch richtigerweise hätte tun sollen. Auch wenn wir in konkreten Fragen unterschiedlicher Meinung sein mögen, wem wir aus welchen Gründen Vertrauenswürdigkeit zuschreiben, gilt dennoch: Diese Zuschreibung bemisst sich daran, ob wir denken, dass die Person das tat, was wir insgesamt als normativ Geboten erachten.

Der Vorschlag hat so formuliert im Übrigen den Vorzug, dass er keine zu enge Konzeption von Vertrauenswürdigkeit darstellt. Vielmehr bleiben mit ihm einige Punkte offen. Auf zwei davon wollen wir kurz hinweisen. Erstens, so wie wir oben mit Thompson festgehalten haben, dass monadische Normen unterschiedlicher Natur sein können (deontisch, auf Etikette bezogen, in der

180 Dieser Kritikpunkt ließe sich auch auf andere Ansätze anwenden, die Vertrauen als die Erwartung verstehen, dass der Vertrauenspartner responsiv auf die eigene Verletzlichkeit ihm gegenüber reagieren (etwa Faulkner 2007, 2014; McGeer 2008; Nickel 2012; Pettit 1995). Diese übersehen (oder machen zumindest nicht explizit), dass Vertrauen und Vertrauenswürdigkeit im Kontext *angemessener normativer* Erwartungen ihren Ort finden.

freundschaftlichen Beziehung zur anderen Person begründet etc.), so bleibt auch unser Ansatz zur Natur der normativen Anforderung an den Vertrauenspartner offen: Vertrauenswürdigkeit kann sich prinzipiell an verschiedenen Arten normativer Anforderung erweisen. Die Vertrauenswürdigkeit eines Arztes etwa zeigt sich darin, dass er die normativen Erwartungen seiner Patienten ihm gegenüber anerkennt und als Maßstab seines Handelns akzeptiert. Er anerkennt diese normativen Erwartungen aber, da es sich um gerechtfertigte *deontische* Erwartungen ihm gegenüber handelt, denen er qua seiner Profession als solcher untersteht. Anders kann sich Vertrauenswürdigkeit in freundschaftlichen Beziehungen darstellen. Bei der normativen Erwartung eines Freundes mir gegenüber, etwa seinen Geburtstag nicht zu vergessen, handelt es sich um eine nicht-deontisch begründete Erwartung.

Der Vorschlag lässt zudem offen, ob die normative Anforderung an den Vertrauensnehmer aus der Natur der Beziehung zum konkreten Vertrauensgeber erwächst oder nicht. Im Beispiel der zwei Freunde ist davon auszugehen, dass einer die normative Erwartung des anderen deshalb ernst nimmt, da ihm die *konkrete andere* Person und die Beziehung zu ihr besonders wichtig ist. Hier kann die Vertrauenswürdigkeit durchaus mit einem persönlichen Wohlwollen der anderen Person gegenüber einhergehen. Im Arzt-Beispiel hingegen kann man davon ausgehen, dass er die normative Erwartung eines Patienten nicht insofern akzeptiert, als es sich um diesen konkreten Patienten handelt. Vielmehr ist er qua seiner Profession darauf verpflichtet (angemessene) normative Anforderungen an ihn von Patienten *überhaupt* Ernst zu nehmen. Trotz dieser Unterschiede gilt für beide Fälle, dass sich Vertrauenswürdigkeit im Kern in der Akzeptanz dessen, was normativ von einem erwartet wird darstellt.

Zuletzt sollen noch zwei Bemerkungen gemacht werden, um mögliche Missverständnisse in Bezug auf den Vorschlag zu vermeiden. Der erste Hinweis ist uns in ähnlicher Form bereits bekannt: Dass die Vertrauenswürdige sich an den normativen Anforderungen ihr gegenüber orientiert, schließt nämlich nicht aus, dass sie *zudem* auch eigene Interessen mit dem Erfüllen des Vertrauens verfolgen kann – beispielsweise die langfristigen Vorzüge einer wechselseitigen Vertrauensbeziehung. Es gilt aber, dass wenn das Verfolgen von Eigeninteressen im Widerspruch damit steht, das zu tun, was normativ geboten ist, sich Vertrauenswürdigkeit einer Person darin zeigt, dass sie von ihren Eigeninteressen ablässt.

Der zweite Hinweis bezieht sich auf einen Strang in der Literatur, der versucht Vertrauenswürdigkeit als Rücksichtnahme auf die *Bedürfnisse* des Vertrauenspartners

auszubuchstabieren.[181] Auch dies muss wieder dahingehend präzisiert werden, dass es nie nur um Bedürfnisse oder Abhängigkeiten *als solche* geht. Vielmehr geht es um diese wiederum nur im Zusammenhang einer damit verbundenen normativen Erwartung an den Vertrauenspartner. Bedürfnisse des Vertrauenden stellen häufig nur den Hintergrund dar, im Lichte dessen bestimmt werden kann, was genau vom Vertrauensnehmer normativ gefordert ist. Angenommen mein Mitbewohner fährt in den Urlaub und bittet mich, seine Pflanzen mit Wasser aus der Regentonne zu gießen und ich verspreche ihm dies zu tun. Im Laufe der Zeit stelle ich jedoch fest, dass das Wasser aus der Tonne unverträglich für Pflanzen ist. In diesem Fall würde ich mich wohl gerade nicht als vertrauenswürdig erweisen, wenn ich mich rigoristisch an die Vorgabe halte und den Pflanzen weiterhin das schädliche Wasser gebe. Besser wäre es, ich würde den Pflanzen stattdessen das Wasser aus der Leitung geben. Auch wenn ich in diesem Fall von dem abweiche, was ich wortwörtlich zugesagt habe, erweise ich mich damit wohl als vertrauenswürdiger.[182] [183] Was normativ von mir erwartet ist, und an welchem Verhalten sich damit meine Vertrauenswürdigkeit erweist, übersteigt in diesem Fall das, was explizit ausgehandelt wurde, und bestimmt sich vielmehr dadurch, was sinnvoller Weise als Bedürfnis des Vertrauenden angenommen werden kann.

5.3.3 Vertrauenswürdigkeit als Tugend

Aufbauend auf diesem Verständnis der Motive, die die Vertrauenswürdigkeit einer Person

181 Vgl. Faulkner 2017, 122; Hartman 2011, 178; Hinchman 2017, 54.

182 Hier zeigt sich ein weiterer Unterschied zwischen vertrauenswürdigem Verhalten und bloßer Pflichterfüllung. Was für vertrauenswürdiges Verhalten erforderlich ist, geht zuweilen über das hinaus, was die Pflicht von einem verlangt. Ähnlich weist Bernard Williams darauf hin, dass sich die Ehrlichkeit eines Sprechers nicht allein dadurch zeigt, dass er wortwörtlich die Wahrheit sagt. So kann sich ein Sprecher auch durch irreführende, wenn auch nicht im strengen Sinne falsche Aussagen, als unehrlich erweisen. Siehe seine Diskussion hierzu in: Williams 2002, 93-122.

183 Dieser Punkt kann auch eine Erklärung dafür geben, warum viele Autor:innen davon ausgehen, dass Vertrauen paradigmatisch mit einer gewissen Diskretion gegenüber dem Vertrauenspartner zusammengeht. Diese Autor:innen prononcieren, dass sich Vertrauen dadurch auszeichnet, dass der Vertrauensperson gewisse Freiheitsgrade zum Nachkommen dieses Vertrauens zugestanden werden (beispielsweise: Baier 1986, 237f; Domenicucci/Holton 2017, 151; Gambetta 1988, 218; Harding 2009, 246; Hartmann 2011, 235-238). David Vitale (2018) hebt dies gar als eines von drei wesentlichen Merkmalen von Vertrauen hervor. Die Betonung dieses Aspektes hat vielleicht auch etwas mit den Arten von Fällen zu tun, an die man bei Vertrauen denkt. Hat man etwa den im Diskurs paradigmatisch herangezogenen Fall der Eltern, die ihrem Babysitter vertrauen vor Augen, verleitet dies dazu, Vertrauen als diskret zu verstehen. Tatsächlich leuchtet mir aber nicht ganz ein, weshalb dies als *wesentlich* für Vertrauen angesehen werden sollte. So gibt es doch auch Fälle, in denen relativ eindeutig ist, was der Vertrauenspartner zu tun hat, um das Vertrauen in ihn nicht zu enttäuschen, etwa wenn ich meinem Mitbewohner vertraue daran zu denken, den Müll herauszubringen.

ausmachen, werden wir im Folgenden die weitergehende These entwickeln, dass Vertrauenswürdigkeit eine Tugend ist. Dabei definiere ich Tugenden gemäß einer aristotelisch angelehnten Minimaldefinition im Wesentlichen als eine feste Disposition (1) dazu, aus guten Motiven heraus zu handeln (2).[184] Sicherlich ließe sich noch detaillierter aufzeigen, inwiefern Vertrauenswürdigkeit einer aristotelischen praktischen Tugend entspricht.[185] Für unsere Zwecke jedoch, beschränken wir uns auf einen Nachweis in groben Umrissen, dass Vertrauenswürdigkeit unter die Kategorie der Tugenden fällt.

ad2) Wir haben oben bereits angedeutet, dass die vertrauenswürdige Person aus guten Motiven handelt. Schließlich handelt sie nicht primär aus Eigeninteresse, sondern aus Einsicht in das, was normativ von ihr verlangt ist. Sie erkennt also die Erfüllung dieser normativen Anforderungen als eine intrinsisch gute Handlung an und folgt ihr aus diesem Grund. Dies kennzeichnet ihr Handlungsmotiv als ein gutes und somit tugendhaftes Motiv.

Aufgrund ihres Zusammenhangs mit der Güte der Handlung ergibt sich im Übrigen auch, wie Aristoteles wiederholt feststellt, dass Tugenden als *lobenswerte* Eigenschaften einer Person zu verstehen sind. Auch dies steht im Einklang mit unserer obigen Bemerkung, dass wir die Vertrauenswürdigkeit einer Person als eine schätzenswerte Eigenschaft dieser Person auffassen.

ad1) Für eine vollständige Theorie der Vertrauenswürdigkeit reicht es allerdings nicht aus deren charakteristische Handlungsmotive zu beschreiben. So ist hier zunächst darauf hinzuweisen, dass eine Person erst insofern als vertrauenswürdig gilt, als sie zumindest auch eine minimale Bereitschaft dazu zeigt, ihre Motive ernsthaft zu verfolgen. In gewissem Sinne reicht es nicht aus, nur guten Willens zu sein oder das „Herz am rechten Fleck“ zu haben. Die wirklich vertrauenswürdige Person ist auch hinreichend dazu *gewillt,* eine gewisse Anstrengungen auf sich zu nehmen.

Dass Vertrauenswürdigkeit auch ausreichende Willensstärke voraussetzt, spiegelt sich auch darin wider, dass es zwei verschiedene idealtypische Formen gibt sich als *nicht* vertrauenswürdig zu erweisen. So kann eine Person mangelnde Vertrauenswürdigkeit zur Schau stellen, indem sie ignoriert, was normativ in einer Situation richtig zu tun wäre, oder sie dem gar bewusst entgegen handelt. Denken wir hierbei etwa an eine Person, die bewusst täuschend agiert. Demgegenüber kann sich aber jemand auch insofern als nicht vertrauenswürdig erweisen, als er zwar schon die

184Siehe: NE II.3 1105a30-33. Siehe auch die Interpretationen in: Annas 2006; Halbig 2013, 84.
185Siehe dazu insbesondere die Überlegungen von: Potter 2002; Kaminski 2020, 209-286.

normative Erwartung ihm gegenüber anerkennt, aber diese nicht mit ausreichender Ernsthaftigkeit zu erfüllen versucht. Wenn ich das Versprechen dir gegenüber, dich im Café zu treffen, deswegen nicht einhalte, weil es regnet und ich zu träge bin das Haus zu verlassen, dann handle ich nicht wirklich täuschend aber nachlässig. Wir können dies auch einen akratischen Fall von Vertrauensverletzung nennen.[186]

Für Vertrauenswürdigkeit heißt dies also, dass sie nicht nur darin besteht, dass man die an einen gestellten normativen Anforderungen als solche anerkennt. Dies muss zugleich auch mit einem gewissen Maß an Entschlossenheit geschehen. An dieser Stelle müssen wir nun aber noch einen Schritt weiter gehen. Um als vertrauenswürdig zu gelten, reicht es nämlich auch nicht, wenn eine Person *sporadisch* eine solche Bereitschaftshaltung zeigt. Wir werden kaum eine Person als vertrauenswürdig bezeichnen, die in dieser Hinsicht volatil ist, d.h. die sich nur gelegentlich bemüht, vertrauenswürdig zu handeln.

Dies zeigt, dass wir Vertrauenswürdigkeit als relativ robuste Disposition einer Person verstehen. Damit ist ein weiteres zentrales Merkmal der aristotelischen Tugendkonzeption erfasst. Um tugendhaft zu sein, reicht es nämlich nicht aus, dass eine Person hin und wieder aus guten Motiven handelt. Sie muss dies auch aus einem gefestigten Charakter heraus zuverlässig tun. Der tugendhafte Mensch befindet sich, wie Aristoteles schreibt, „fest und ohne Wanken“ in einem Zustand[187], der ihn dazu befähigt, die richtigen Handlungsgründe zu erkennen und entsprechend aktiv zu werden.

Hierbei soll das Prädikat „robust“ genau genommen zwei Aspekte festhalten.[188] Erstens zählt eine Tugend insofern als robust, als sie in diachroner Perspektive stabil ist. Eine tugendhafte Person ist nicht spontan, von einem Tag auf den anderen, nicht mehr tugendhaft. Zweitens zeigt sich die Robustheit der Tugenden darin, dass sich die tugendhafte Person über verschiedene Handlungskontexte hinweg invariant tugendhaft verhält. Insofern stellt sich die Tugendhaftigkeit einer Person ja gerade in solchen Situationen unter Beweis, in denen es herausfordernd ist,

186Der Unterschied zwischen beiden Fällen scheint sich auch anhand zweier möglicher Formulierungen des Vorwurfes im Falle verletzten Vertrauens zu spiegeln: „Du hast mein Vertrauen enttäuscht“ einerseits, und „Du hast mein Vertrauen missbraucht“ andererseits. Der letztgenannte Vorwurf scheint tendenziell dann angebracht, wenn die andere Person eine täuschende Absicht gezeigt hat. Allgemein weist der Begriff „Missbrauch“ darauf hin, dass etwas nicht für oder womöglich sogar entgegen seinem eigentlichen Zweck verwendet wird (vgl. hierzu die Überlegungen von O‘Neill 2012). Vertrauen zielt, wie wir oben gesehen haben, wesentlich auf eine kooperative Interaktionsform ab. Wer das Vertrauen nicht nur vernachlässigt, sondern es nur akzeptiert, um es als Mittel für andere Zweck zu verwenden, der missbraucht dieses. Demgegenüber zeigt der Vorwurf „Du hast mein Vertrauen enttäuscht“ tendenziell an, dass die Vertrauensperson zwar prinzipiell gute Absichten hatte, diesen aber nicht voll und ganz gerecht geworden ist.

187NE II.3 1105a33.

188Siehe Halbig 2013, 96-101.

tugendhaft zu handeln. Hierin zeigt sich im Übrigen auch, dass die tugendethische Konzeption das obige Problem der fehlenden Stabilität, das gegen die absichtsbezogenen Ansätze sprach, umgeht.

Dieser Punkt bietet im Übrigen auch eine Erklärung für eine weitere begriffliche Eigenschaft von Vertrauenswürdigkeit, nämlich die, dass es sich bei Vertrauenswürdigkeit offenbar um ein graduelles Konzept handelt. Das soll heißen, dass wir Personen offensichtlich als *mehr oder weniger* vertrauenswürdig beschreiben können.[189] Dies ließe sich so verstehen, dass wir Personen als *mehr oder weniger* robust in Bezug auf die an sie gestellten normativen Anforderungen begreifen.

Zusammenfassend lässt sich sagen, dass zwei Merkmale von Vertrauenswürdigkeit dafür sprechen, diese als Tugend aufzufassen: Zum einen handelt die vertrauenswürdige Person auf der Ebene der Motive aus guten und lobenswerten Motiven. Zum anderen demonstriert sie eine stabile Bereitschaft dazu aus diesen Motiven zu handeln. Alles in allem erscheint Vertrauenswürdigkeit somit als eine feste Disposition, um der Güte der beabsichtigten Handlungen willen zu handeln.

THESE

Vertrauenswürdigkeit stellt eine Tugend dar, d.h. sie ist eine:

1. **feste Disposition einer Person dazu**
2. **aus guten Motiven zu handeln.**

5.3.4 Vertrauenswürdigkeit und Kompetenzen

Die tugendtheoretische Konzeption ist nicht mit der zuweilen in Literatur geäußerten These zu verwechseln, dass Vertrauenswürdigkeit mit den für den konkreten Fall geeigneten Kompetenzen einhergehen muss.[190] Auch in vielen sozialwissenschaftlichen Ansätzen wird Kompetenz gerne als eine (neben mehreren) Komponente von Vertrauenswürdigkeit dargestellt. Vertrauenswürdigkeit wird dann etwa verstanden als Summe von Kompetenz und moralisch guten Motiven.[191] Insofern Kompetenzen dabei aber wirklich nur im Sinne von Fähigkeiten oder Know-How gesehen werden

189Zuweilen sprechen wir natürlich auch in binärer Form und stellen fest: „Person P ist vertrauenswürdig“ oder „Person Q ist nicht vertrauenswürdig“. In dieser Redeweise ließen sich die Begriffe „vertrauenswürdig/nicht vertrauenswürdig“ als *Schwellenbegriff* verstehen (vgl. Halbig 2013, 103). Das heißt, wir beschreiben eine Person dann als vertrauenswürdig *simpliciter*, sofern ihre Vertrauenswürdigkeit eine bestimmte Schwelle überstiegen hat.

190Wie etwa von Karen Jones nahelegt (Jones 1996, 1).

191Vgl. zum Beispiel: Johnson 1999; Mayer et al. 1995.

(zuweilen ist der Begriff aber auch nicht immer klar definiert) sind sie von Tugenden abzugrenzen. Im Gegensatz zu Tugenden gehen diese nämlich nicht mit einem Gespür für das normativ erforderliche einer Situation, geschweige denn mit der Bereitschaft einher entsprechend zu handeln. Ich kann etwa die Fähigkeit haben Klavier zu spielen. Das bedeutet erstens noch nicht, dass ich weiß, wann es angemessen oder gar geboten ist Klavier zu spielen, und zweitens bedeutet es noch weniger, dass ich nur dann spiele, wenn ich spielen darf bzw. sollte.

Gemäß unserer tugendethischen Konzeption haben demnach Kompetenzen, als bloße Fähigkeiten verstanden, als solche nichts mit Vertrauenswürdigkeit zu tun. Tatsächlich scheint dies aber auch eine plausible Sichtweise zu sein, wohingegen die Komponententheorie die eigentlichen begrifflichen Zusammenhänge verzerrt wiederzugeben scheint. Es sind schließlich niemals mangelnde Kompetenzen *als solche*, die eine Person bereits weniger vertrauenswürdig machen. Spiegelbildlich muss Zweifel an der Kompetenz einer anderen Person als solches noch nicht zwangsläufig Misstrauen darstellen. Nur unter bestimmten Bedingungen, die ich gleich erläutern werde, drückt mangelnde Kompetenz auch mangelnde Vertrauenswürdigkeit aus.

Besonders einleuchtend ist dieser Gedanke in Bezug auf *commitment*-basiertem Vertrauen. Ganz grundlegend zeigt sich für diese eine besondere normative Anforderung, die eine vertrauenswürdige Person zu berücksichtigen weiß. Wie Hawley herausstellt, geht die vertrauenswürdige Person nämlich nur in dem Maße *commitments* ein, wie sie dazu in der Lage ist diese auch einzuhalten.[192] Ein Korollar dieser Aussage ist, dass eine genuin vertrauenswürdige Person nur dort zu Vertrauen einlädt, wo sie auch die Kompetenzen hat diesem Vertrauen nachzukommen. Indem man nämlich einem Vertrauenspartner ein bestimmtes *commitment* gegenüber macht (etwa ϕ zu tun), kommuniziert man damit in der Regel auch, dass man die geeignete Kompetenz dazu habe, das *commitment* zu erfüllen (ϕ zu tun).[193] Wenn ich meinem Freund verspreche eine Treppe in seinem zu konstruieren, gebe ich ihm damit implizit auch zu verstehen, dass ich mich als fähig genug für dieses Unterfangen ansehe. Vertrauenswürdigkeit besteht dann aber nicht in den Kompetenzen *als solchen*, sondern in der ernsthaften Absicht *nur dort zu Vertrauen einzuladen, wo man die geeigneten Kompetenzen auch hat*.[194] An den Kompetenzen *erweist* sich also mitunter, ob jemand vertrauenswürdig ist. Dies ist aber nicht das

192Hawley 2019, 82.

193An diesem Merkmal orientiert sich auch Jones' Begriff der *reichen Vertrauenswürdigkeit*: „*B is richly trustworthy* with respect to A just in case: (i) B is willing and able reliably to signal to A those domains in which B is competent and will take the fact that A is counting on her, were A to do so, to be a compelling reason for acting as counted on and (ii) There is a non-trivial number of relatively central domains in which B will be responsive to the fact of A's dependencyin the manner specified in (i)" (Jones 2012b, 74).

194Für eine ausführlichere Diskussion diesbezüglich, siehe wiederum Hawley 2019, 72-142.

Gleiche wie zu behaupten, dass sich die Vertrauenswürdigkeit durch die Kompetenz *konstituiert*, wie die erwähnten Komponententheorien nahelegen.

Zusammenfassend besagt das Ergebnis dieser Diskussion, dass Vertrauenswürdigkeit eine verlässliche Disposition dazu darstellt, sich intrinsisch und ernsthaft in seinem Handeln an dem zu orientieren, was normativ Geboten ist. An dieser Charakterisierung werden wir im weiteren Verlauf der Arbeit so festhalten. Ergänzungsbedürftig ist hingegen weiterhin unsere bisherige Konzeption interpersonellen Vertrauens selber, wir wir im folgenden Kapitel genauer aufzeigen werden.

6. Vertrauen als Anerkennung

Blicken wir zurück auf das, was wir bisher über interpersonelles Vertrauen festgehalten haben, dann lässt sich das Ergebnis folgendermaßen zusammenfassen. Interpersonelles Vertrauen stellt eine risikobereite Einstellung dar, für die im Wesentlichen zwei Elemente charakteristisch sind: Einerseits eine *normative Erwartung* an den Vertrauenspartner vertrauenswürdig zu handeln, und andererseits *die Annahme*, dass der Vertrauenspartner auch tatsächlich vertrauenswürdig ist.

Tatsächlich kann man feststellen, dass dies eine in der Vertrauensliteratur ganz standardmäßig vertretene Analyse darstellt. Im Folgenden werden wir diesbezüglich auch von der *Standardauffassung interpersonellen Vertrauens* sprechen. Ein Paradebeispiel derselben findet sich etwa bei Margaret Urban Walker. Ihre Vertrauenskonzeption „captures two elements shared by a wide variety of cases of trust: expectation of others to perform as relied upon, and the "participant attitude" toward reliance in which I am prepared to hold you responsible for doing what I assume you *should*."[195] Im Einzelnen mag diese Auffassung verschiedene nuancierte Variationen finden. Diese mögen sich in der konkreteren Ausbuchstabierung des Ansatzes unterscheiden – etwa in Bezug darauf, was genau unter Vertrauenswürdigkeit zu verstehen ist, ob die Annahme, dass der Vertrauenspartner vertrauenswürdig ist, hier eine epistemische Erwartung oder Überzeugung darstellt, oder was die Rechtfertigung der normativen Erwartung angeht. Was sie aber über alle Varianz hinweg auf einer grundlegenderen Ebene teilen, ist ein bestimmtes zweiteiliges

195 Walker 2006, 80. Auch wenn Walker das grobe Schema unserer Analyse teilt, weicht ihr Ansatz in den Details von unserer obigen Position ab. Dies betrifft insbesondere die Auffassung von Vertrauenswürdigkeit. Näher ist uns da schon Paul Faulkners Definition interpersonellen (in seinen Begriffen „affektiven") Vertrauens, die ebenso normative Erwartungen und einen epistemischen Optimismus zusammenbringt: „*A* trusts *S* to *Φ* (in the affective sense) if and only if (1) *A* knowingly depends on *S Φ*-ing and (2) *A* expects *S*'s knowing that *A* depends on *S Φ*-ing to motivate *S* to *Φ* (where *A* expects this in the sense that he expects it of *S*)." (Faulkner 2011, 146).

Analyseschema, das ich folgendermaßen zusammenfassen möchte:

Die Struktur der Standardauffassung von Vertrauen:

Einer Person bezüglich Φ zu vertrauen, beudetet, ihr gegenüber in Bezug auf Φ ein Risiko einzugehen, und:

1. eine normative Erwartung an P zu haben, vertrauenswürding zu handeln
2. zuversichtlich zu sein, dass P vertrauenswürdig ist

Im Folgenden werden wir ungeachtet der Details seiner verschiedenen Versionen dieses allgemein geteilte Schema als solches diskutieren. Oben haben wir vereinzelt bereits angedeutet, dass der Begriff interpersonellen Vertrauens mit diesem noch nicht erschöpfend geklärt ist. Diesen Verdacht weiter zu erhärten und klarer zu formulieren, wird das Ziel des vorliegenden Kapitels sein. Genauer gesagt werden wir dies in einem ersten Schritt anhand einer Diskussion des Problems nennen, das ich als Problem der Sozialkalkulation bezeichne (6.1). Demnach gelingt es der Standardauffassung nicht wirklich zufriedenstellend Vertrauen von einer bloß sozialkalkulatorischen Einstellung abzugrenzen.

Vor diesem Hintergrund werden wir präziser die theoretische Herausforderung für den Rest dieses Kapitels artikulieren können. Diese besteht darin, einen Ansatz auszuarbeiten, der Vertrauen als eine im Vergleich zur sozialkalkulierenden Einstellung *persönlichere* Einstellung gegenüber dem Vertrauenspartner verständlich macht. Die Idee, die wir anschließend entwickeln, besteht darin, Vertrauen als eine *anerkennende* Haltung zu begreifen. Diesen Gedanken werden wir maßgeblich in kritischer Auseinandersetzung mit dem Beitrag von Richard Moran zur Zeugenschaftsdebatte entwickeln (6.2/6.3). Darüber hinaus gibt es aber auch eine Reihe weiterer Überlegungen, die ein Verständnis von Vertrauen als Anerkennung nahelegen, sodass der anerkennungsbezogene Ansatz ein insgesamt kohärentes und phänomenal adäquates Bild interpersonellen Vertrauens zu zeichnen verspricht. Dies betrifft etwa auch die Darlegung des begrifflichen Zusammenhanges zwischen Vertrauen und der Beziehung zwischen den beiden Vertrauenspartnern, die mithilfe des Anerkennungsansatzes vorgenommen werden kann (6.4). Nicht zuletzt wird dabei auch ein maßgebliches Ziel sein, aufzuzeigen, dass unsere Vertrauenskonzeption weit genug ist, um der breiten Vielfalt an möglichen Vertrauensbeziehungen gerecht zu werden.

6.1 Das Problem der Sozialkalkulation

Um in diesem Kapitel die problematische Limitierung der Standardauffassung in einem schrittweise entfalteten Argumentationsgang aufzuzeigen, werden wir zunächst einmal darlegen, dass es zur Charakterisierung interpersonellen Vertrauens nicht ausreicht zu postulieren, dass Vertrauen mit einer normativen Erwartung an den Vertrauenspartner einhergeht, wie es die erste Prämisse der Standardauffassung festhält. Wie wir aufzeigen werden, erscheint die normative Erwartung nur als begleitendes, aber nicht als konstitutives Moment von Vertrauen. Dazu sei hier noch einmal das Anscombe'sche Beispiels des gescheiterten Lügners aufgerufen. Angenommen eine Person habe vor mich bezüglich der Frage, ob p oder non-p der Fall ist zu belügen. Stellen wir uns außerdem vor, ich wüsste nicht nur von dieser Täuschungsabsicht des Sprechers, sondern auch davon, dass er bezüglich p falsch informiert ist. Wenn er mir nun also mitteilte, dass p der Fall ist, dann kann ich gewissermaßen über zwei Ecken darauf schließen, dass dies der Wahrheit entspricht.

Der Fall erfüllt damit nicht nur die Kriterien für Vertrauen im generischen Sinne. Immerhin gehe ich ein Risiko damit ein zu glauben, dass p zutrifft, bin aber zugleich auch zuversichtlich diesbezüglich. Darüber hinaus mag ich aber auch eine normative Erwartung gegenüber dem Sprecher haben: So wie man allgemein die Wahrheit sprechen sollte, sollte auch er es tun.

Dennoch gilt für diesen Fall wohl kaum, dass ich *dem Sprecher* vertraue – immerhin weiß ich, dass mich dieser belügen wollte! Es zeigt sich damit, dass sich an dem bloßen Umstand, dass ich eine normative Erwartung einem Sprecher gegenüber einnehme, in keinster Weise entscheidet, ob ich ihm vertraue oder nicht. Diese normative Erwartung – wir sprachen diesen Punkt in der Diskussion um den normativen Kontext von Vertrauen an – besteht schließlich auch *unabhängig* von meinem Vertrauen oder Misstrauen. Allein *qua Sprechakt* und unabhängig davon, ob ich ihm diesen abkaufe oder nicht, hat sich der Sprecher auf die Wahrheit des Gesagten verpflichtet – ein Gedanke, den etwa auch Richard Moran formuliert, auf dessen Beitrag wir unten noch genauer eingehen werden: „the speaker, in presenting his utterance as an *assertion*, one with the force of *telling* the audience something, presents himself as *accountable* for the truth of what he says".[196] Der Sprecher berechtigt mich qua seines Sprechaktes also einen entsprechenden normativen Anspruch ihm gegenüber zu haben, mit den entsprechenden Implikationen, wie etwa der, dass ich dazu berechtigt wäre ihm Vorwürfe zu machen, falls er gelogen hätte. Dies gilt aber vollkommen

196Moran 2005, 11. Dies muss nicht heißen, dass einem Sprecher in jedem Fall Vorwürfe gemacht werden können, wenn das, was er sagt, nicht der Wahrheit entspricht. Immerhin können etwa entschuldigende Umstände vorliegen (der Sprecher konnte etwa nicht erwartbar wissen, dass er sich irrte). Dennoch bleibt es prinzipiell dabei, dass er als Sprecher wesentlich normativ an der Wahrheit des Gesagten gemessen wird (Vgl. Faulkner 2011, 140f.).

unabhängig davon, ob ich diesbezüglich Vertrauen in ihm habe oder nicht.

Vertrauen findet also zwar stets seinen angemessenen Ort in Bezug auf normative Erwartungen, umgekehrt aber bedeutet die Berechtigung zu einer solchen normativen Erwartung *als solches noch nicht*, dass man der anderen Person vertraut. Es liegt eher nahe, dass die normative Erwartung nur ein *begleitendes Element* von Vertrauen sein kann: Immerhin konstituiert der normative Kontext zwar erst den Möglichkeitsraum für Vertrauen (und damit auch Misstrauen). Ob diese Möglichkeit *realisiert* wird oder nicht, entscheidet sich an einer anderen Stelle. Zumindest ohne weitere Erläuterung leistet das Postulat einer normativen Erwartung im Falle interpersonellen Vertrauens noch keinen substanziellen Beitrag dazu, zu klären, was es mit dem Vertrauen in eine Person auf sich hat.

Es liegt deshalb nahe zu sagen, dass das Vertrauen in eine Person durch ein anderes Moment konstituiert ist. Ein Blick auf das Beispiel scheint hier auch direkt eine uns bereits bekannte Idee nahezulegen: Es scheint, dass wir in dem Beispiel nicht davon sprechen würden, dass ich dem Sprecher vertraue, da ich ihm keine vertrauenswürdige Einstellung unterstelle. Schließlich wollte er mich ja eigentlich belügen. Dies begründet die Vermutung, dass sich Vertrauen im Kern dadurch auszeichnet, dass man sich auf die Vertrauenswürdigkeit des Vertrauenspartners bezieht.

Diese Idee ist auch in der zweiten Prämisse der Standardauffassung eingefangen. Allerdings zeigt sich, dass auch mit dieser die Standardauffassung zu unpräzise ist. Nach unserer Darstellung formuliert sie lediglich, dass die vertrauende Person *annimmt*, dass die andere Person vertrauenswürdig ist. Dies jedoch ist eine zu weit gefasste Charakterisierung, da sie in problematischer Weise einen bloßen sozialkalkulatorischen Zugriff auf die Vertrauenswürdigkeit der anderen Person nicht ausschließt. Um dieses Problem genauer darzulegen, betrachten wir zunächst das folgende prominente *Heiratsschwindler-Beispiel*:[197]

> Der überzeugende Heiratsschwindler Adam bringt sein Opfer Berta dazu, ihm zu versprechen ihn zu heiraten. Er hat dabei Berta gut studiert und weiß, wie sehr diese gewillt ist, sich an ihn zu binden. Deswegen ist Adam auch ziemlich zuversichtlich, dass sein Plan aufgehen und Berta sich an ihr Versprechen halten wird.

Auch dieser Fall erfüllt zunächst die Kriterien für generisches Vertrauen: Der

197Das Beispiel wurde meines Wissens erstmals von Holton (1994) angeführt und seitdem in der Diskussion um Vertrauen vereinzelt diskutiert, beispielsweise in: Cogley 2012; Hawley 2014a; Lahno 2002.

Heiratsschwindler geht mit dem was er tut sicherlich ein Risiko ein, ist aber dennoch zuversichtlich. Bemerkenswert ist aber, dass sich seine Zuversicht gerade darauf stützt, dass er von der Vertrauenswürdigkeit seines Opfers überzeugt ist. Die zweite Prämisse der Standardauffassung ist hier also erfüllt. Und dennoch, würden wir kaum davon sprechen, dass der Heiratsschwindler *seinem Opfer* vertraue.[198]

Noch eindrücklicher wird dieser Punkt, wenn wir uns auf ein Beispiel aus dem Bereich der Zeugenschaft beziehen. Betrachten wir zur zusätzlichen Veranschaulichung etwa folgenden Fall:

> Mein Mitbewohner teilt mir mit, dass er im Keller weiße Mäuse gesehen hat. Stellen wir uns vor, ich stellte daraufhin ein Kalkül der folgenden Art an: Ich vergegenwärtige mir all die Fälle, in denen mein Mitbewohner in der Vergangenheit etwas Vergleichbares gesagt hat und überprüfe ob dies jeweils der Wahrheit entsprach. Man kann sich vielleicht auch vorstellen, dass ich gewissermaßen Buch darüber geführt hätte und feststelle, dass seine Aussagen auf zuverlässige Weise mit der Wahrheit des Gesagten koinzidieren. Ich prüfe zudem, ob vielleicht noch besondere Bedingungen vorliegen, die fragwürdig machen, ob er die Wahrheit sagt, ob er etwa unter optischen Illusionen leidet, ob er ein Interesse daran hätte mich glauben zu machen, dass im Keller weiße Mäuse sind etc. etc. Auf Grundlage dieser indizienbezogenen Kalkulation kann ich nun zu der Schlussfolgerung gelangen, dass mein Mitbewohner in ausreichendem Maße vertrauenswürdig ist und, dass das was er sagt höchstwahrscheinlich der Wahrheit entspricht.

Auch hier scheint es nicht angemessen zu sagen, dass ich meinem Mitbewohner vertraue. Man beachte schließlich, dass ich in dem Beispiel allein ein indizienbezogenes Kalkül anstelle, dass mich darauf schließen lässt, dass mein Mitbewohner vertrauenswürdig ist, und in der Folge auch, dass er die Wahrheit sagt. Was ich damit jedoch nicht mache, ist, *meinem Mitbewohner* zu vertrauen. Ich schließe allein auf das *Faktum*, dass mein Mitbewohner vertrauenswürdig ist, und von dort aus auf die Wahrheit dessen was er sagt. Lars Hertzberg drückt in einem ähnlichen Zusammenhang diesen Gedanken einmal sehr prägnant aus: „When I trust someone, it is *him* I trust; I do *not trust certain things about him*.“[199] Wobei unter „certain things“ im Zusammenhang

198 Vielleicht mag man einwenden, dass es sich dabei um keine besonders realistische oder psychologisch akkurate Beschreibung eines typischen Heiratsschwindlers handelt. Das mag sein. Der Punkt ist an dieser Stelle jedoch ein begrifflicher, der mit Hilfe des fiktiven Heiratsschwinderfalls illustriert werden soll. Die Standarddefinition lässt zu, dass „Vertrauen“ als bloße Sozialkalkulation verstanden werden könnte, für die der Heiratsschwindler nur eine anschauliche Verkörperung darstellt.

199 Hertzberg 1988, 315; Hervorhebung von mir. Hertzberg jedoch zieht den Unterschied zwischen Sich-Verlassen und

unseres Beispiels der Umstand zu verstehen wäre, dass mein Mitbewohner vertrauenswürdig ist. Darauf zu vertrauen, dass dies der Fall ist, ist aber etwas anderes als *ihm* zu vertrauen. Und dieser Umstand, dass ich meinem Mitbewohner nicht vertraue, wird auch nicht ohne Weiteres dadurch aufgehoben, dass ich außerdem eine normative Erwartung ihm gegenüber habe. Ich mag denken, dass mein Mitbewohner die Wahrheit sagen sollte, aber dadurch vertraue ich ihm noch nicht.

Dieser Fall lässt sich auch mit unter dem Vorzeichen der Leitunterscheidung aus dem Zeugenschaftsdiskurs diskutieren. In dieser wird üblicherweise ein Unterschied gezogen zwischen dem was es heißt „zu glauben, was jemand sagt“ und „*jemandem* zu glauben, was er sagt“[200] bzw. „jemandem zu vertrauen“. Zu glauben was jemand sagt bedeutet dabei lediglich, dass man von der Wahrheit des Gesagten überzeugt ist. Allein die Wahrheit des Gesagten zu akzeptieren, bedeutet aber nicht einer Person zu glauben – auch wenn man sich dabei auf Indizien stützt, die für die Vertrauenswürdigkeit einer Person sprechen. Jemandem zu glauben, als spezieller Fall von Vertrauen in jemanden, geht auf eine noch zu klärende Weise darüber hinaus.

Bei genauerem Hinsehen haben wir in dem Zeugenschaftsbeispiel einen strukturell ähnlichen Fall wie in dem vorangegangen *Heiratsschwindler-Beispiel* vorliegen. Die zwei Fälle scheint zu verbinden, dass der jeweils „vertrauende“ Akteur allein einen eigenen Zweck verfolgt (sei es an Geld oder zu einer wahren Überzeugung zu gelangen), für die der Vertrauenspartner ein probates Mittel darstellt.[201] Die Überzeugung, dass der andere vertrauenswürdig ist, kommt in dieser Perspektive der Überzeugung gleich, dass das Mittel zum Erreichen dieses Zweckes auch verlässlich ist. Es spielt hierbei keine Rolle im Lichte welchen Zweckes genau ein Akteur an der Frage interessiert ist, ob die andere Person vertrauenswürdig ist: Sofern sie dabei die Vertrauenswürdigkeit nur als Mittel zum Erreichen dieses Zweckes ansieht, hat dies nichts mit Vertrauen *in die Person* zu tun.

Aus einer anderen Richtung kommend kann man das Problem auch so ausdrücken, dass es der Standardauffassung des Vertrauens bei genauerer Betrachtung nicht gelingt, einen tragbaren Unterschied zwischen Vertrauen und Sich-Verlassen zu ziehen. Vielmehr lässt die Standardauffassung zu, dass Vertrauen auch nur als ein *spezieller Anwendungsfall oder bestimmte Konfiguration* eines bloßen Sich-Verlassen angesehen werden kann. Oben haben wir Sich-Verlassen in Form einer bloßen epistemischen Annahme bezüglich des Verhaltens einer anderen

Vertrauen anders als in dieser Arbeit nicht anhand der Kategorie der Anerkennung.

200 Siehe etwa die Erläuterung dieser Unterscheidung in: Faulkner 2011, 55f.

201 Der Punkt ist hier nicht einmal, dass es sich dabei um ein *eigennütziges* Interesse handeln muss. Vielmehr ist entscheidend, dass es mit Vertrauen in eine Person nicht vereinbar ist, diese als bloßes Mittel für einen Zweck zu betrachten – unabhängig davon, ob es sich um einen persönlichen oder kollektiv geteilten Zweck handelt.

Person definiert, auf deren Grundlage man praktische Vorhaben beschließt, auch wenn man im Zuge dessen ein Risiko eingeht. Die Standardsichtweise scheint nun angesichts des Problems der Sozialkalkulation immer noch bestimmte Fälle dieser Art als Beispiele von Vertrauen zuzulassen. Dieses Sich-Verlassen wird lediglich als in zwei Punkten *spezifisch* ausgewiesen: Erstens insofern es notwendig mit einer, allerdings ohnehin bestehenden, normativen Erwartung begleitet wird, und zweitens, als sich die epistemische Erwartung nicht einfach nur auf das bloße *Verhalten* der anderen Person bezieht, sondern sich auch auf das vertrauenswürdige Motiv des anderen stützt. Der Heiratsschwindler, so kann man entsprechend sagen, vertraut nicht der anderen Person, er *verlässt sich* lediglich auf ihre Vertrauenswürdigkeit.

Zu diesem Problem der Sozialkalkulation muss noch eine wichtige methodische Bemerkung gemacht werden. Vielleicht mag man einwenden, dass die bemühten Beispiele eine psychologisch recht unplausible Beschreibung davon liefern, wie Personen innerlich eingestellt sind. Es scheint fragwürdig, ob man wirklich eine derart distanzierte und rein kalkulierende Perspektive anderen Personen gegenüber jemals einnehmen kann. Das mag sein. Die Beispiele hatten allerdings nicht den methodischen Zweck eine psychologische akkurate Beschreibung abzugeben, sondern vielmehr einen begrifflichen Punkt zu illustrieren. Folgendes ist zu beachten: Insofern man wie in den Beispielen eine Person als bloß sozialkalkulierend eingestellt beschreibt, schreibt man dieser eine bestimmte Perspektive, eine bestimmte Art des Zugriffes auf die andere Person zu – es sei dahingestellt, ob es realiter möglich ist, diese durch und durch einzunehmen. Die Beispiele zeigen nun, dass es im Rahmen dieses sozialkalkulierenden *Ausblickes* auf die andere Person prinzipiell nicht möglich ist, die Bezugnahme zu dieser als personales Vertrauen auszuweisen. In der sozialkalkulierenden Betrachtungsweise gibt es keinen begrifflichen Raum für Vertrauen. Der in der Vertrauensliteratur häufiger diskutierte Fall des Heiratsschwindlers, so schlage ich hier vor, muss demnach wesentlich als Illustration dieses Problem der Begrenzung eines bestimmten kategorialen Zugriffes gesehen werden. Das Problem an der Standardauffassung ist demnach, dass sie eine rein sozialkalkulatorische Ausdeutung ihrer Begriffsbestimmung nicht ausschließt.[202]

Das Problem der Sozialkalkulation, das hier zutage tritt, stellt eine notorische, häufig unbeachtete, Herausforderung in der gegenwärtigen Vertrauensliteratur dar. Demnach lässt ein

202 Peter Strawson meint in ähnlicher Weise, dass es *anthropologisch* nicht möglich sei, allein eine rein objektive Haltung, das heißt nur berechnende Haltung anderen Personen gegenüber einzunehmen. Dennoch hält er an dem Konzept der objektiven Haltung als einer bestimmten *Perspektive*, einem bestimmten Blickwinkel fest, in dem wir andere Personen betrachten können. Ebenso halte ich es auch mit der sozialkalkulierenden Perspektive, die wie wir gleich sehen werden, auch einer solchen objektiven Haltung zuzuordnen ist.

Großteil der vorfindlichen Ansätze zur Definition und Modellierung interpersonellen Vertrauens zu, dass sich Vertrauen auch als eine bloß sozialkalkulatorische Haltung verstehen ließe.[203] Um ein eindrückliches Beispiel für dieses verbreitete Defizit in der Theoriebildung herauszugreifen, können wir etwa den Ansatz von Philipp Pettit betrachten. Auch Pettit versucht genuines Vertrauen von einem bloßen Sich-Verlassen abzugrenzen. Zur Illustration des Spezifikum genuinen Vertrauens bezieht er sich auf das Beispiel des Vertrauens in einen Busfahrer:

> „I see the driver as someone who is well-disposed toward me, whether in my individual right or as a member of the public, and I believe that my manifesting reliance will strengthen or reinforce his existing reasons to do that which I rely on him to do. For whatever reason, I assume the attitude of a trusting individual."[204]

Ohne weitere Qualifizierung aber droht dieser Definitionsversuch nun zuzulassen, dass auch Fälle, wie die beiden oben beschriebenen, als Fälle genuinen Vertrauens einzuordnen sind. Der Heiratsschwindler nimmt seinem Opfer gegenüber immerhin eine vergleichbare Einstellung ein, wie die Person in Pettits Beschreibung gegenüber dem Busfahrer. Beide sind demnach von den guten Motiven der anderen Person überzeugt. Insofern man sich auf diese Motive jedoch bloß als Element in einem instrumentellen Kalkül bezieht – und dies ist es, was Pettits Charakterisierung nicht ausschließt – scheint es, wie gesagt, unangebracht, seine Haltung als Vertrauen zu bezeichnen.

Wie aber ist dann Vertrauen eigentlich zu verstehen? Was genau fehlt an der Haltung des Sozialkalkulierenden, um diese als vertrauende Haltung zu akzeptieren? Bezüglich dieser Frage scheint es in der Vertrauensliteratur eine artikulatorische Leerstelle zu geben. Wie gesagt können die meisten Modellierungen von Vertrauen als eine bloß sozialkalkulierende Haltung verstanden werden. In den Ansätzen, in denen ein Problembewusstsein diesbezüglich durchscheint, treten tendenziell Schwierigkeiten auf *positiv* zu formulieren, wodurch sich Vertrauen auszeichnet.

203 Neben den oben genannten Positionen der Standardauffassung von Vertrauen, stellt sich dieses Problem etwa auch für Annette Baiers Charakterisierung von Vertrauen als „reliance on another's good will" (Baier 1986, 234). Ebenso stellt es sich für Jones' Vorschlag, Vertrauen bestünde in der optimistischen Vorstellung, der Vertrauenspartner werde durch die eigene Abhängigkeit zu ihm „directly und favourably moved" (Jones 1996, 4). Siehe auch die Kritik an beiden in: Cogley 2012. Ein besonders eindrückliches Beispiel stellt Benjamin McMylers zweitpersonale Vertrauenstheorie dar. So definiert McMyler Vertrauen folgendermaßen: „If I trust Mary, then (1) I am thereby recognizing her as having a certain second-personal competence and authority; (2) there is a sense in which she is responsible to me for doing what I trust her to do; and (3) if she does not do what I trust her to do, I will thereby be entitled to adopt secondpersonal attitudes like resentment back towards her" (McMyler 2011, 127). Diese Definition bedürfte sicherlich einer ausführlicheren Diskussion, sie scheint mir aber ohne präzise Qualifizierung, ebenso eine bloß sozialkalkulatorische Deutung nicht auszuschließen.

204 Pettit 1995, 206.

Richard Holton etwa markiert Vertrauen, wie wir gesehen haben, mit Strawons Begriff einer teilnehmenden Haltung (*participatory attitude*).[205] Damit setzt er sie in Kontrast zu dem, was bei Strawson eine objektive Haltung (*objective attitude*) genannt wird. Jene Perspektive zeichnet sich nach Strawson dadurch aus, dass die wohlwollende oder böswillige Einstellung des Anderen *als solche* belanglos ist. Die andere Person wird lediglich in objektivierter Form als etwas in seinem Verhalten zu Verstehendes oder zu Prognostizierendes gesehen. Beispielsweise kann diese Perspektive im Rahmen gewisser instrumenteller Überlegungen, etwa sozialtechnischer Art, eingenommen werden: „To adopt the objective attitude to another human being is to see him, perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided".[206] Die bloß verhaltensprognostizierende Einstellung des Heiratsschwindlers ließe sich hier also reibungslos in die Aufzählung mit einreihen.

Der Begriff der partizipatorischen Haltung kommt damit zwar als attraktiver Gegenbegriff in den Blick. Das Problem ist jedoch, dass dieser Begriff auch von Holton nicht wirklich substanziell erläutert wird. In einer dünnen Charakterisierung macht Holton lediglich darauf aufmerksam, dass eine partizipatorische Haltung sich in der Bereitschaft zeigt reaktive Einstellungen zu manifestieren. Insofern reaktive Einstellungen jedoch wiederum allein als Ausdruck einer normativen Erwartung verstanden werden (die wir oben als lediglich begleitend ausgewiesen hatten, zumindest solange sie nicht näher erläutert werden) hilft dieser vage Hinweis nicht weiter, um das Problem der Sozialkalkulation zu umgehen.[207]

Stephen Darwall auf der anderen Seite macht ebenfalls darauf aufmerksam, dass es zu kurz kommt Vertrauen als mit normativen Erwartungen einhergehend zu beschreiben, die sich in reaktiven Einstellungen wie moralischen Vorwürfen zeigen. Er strebt deshalb an, Vertrauen als „some form of personal relation"[208] auszuweisen. Seine Erläuterungen dazu, was genau diese Form personaler Relation kennzeichnet sind jedoch recht tentativ. Er formuliert: „trust is a species of second-personal attitude through which we lay ourselves open to others in a way that is distinctive of personal relationship and attachment."[209] Darüber hinaus merkt er einige Parallelen zwischen Vertrauen und Liebe an. Dies droht nicht nur ein zu enges Begriffsverständnis von Vertrauen zu

205Holton 1994, 67.
206Strawson 1968, 79.
207Siehe hierzu auch die Diskussion von Zac Cogley (2012, 36-38).
208Darwall 2017, 46.
209Ebd., 46.

forcieren, in Bezug auf das sich die Frage stellt, wie hier formalere oder flüchtigere Formen von Vertrauen noch ihren Platz finden können. Es stellt zudem auch noch keine substanziell ausgearbeitete *Theorie* von Vertrauen dar.

Insgesamt können wir am Ende dieser Diskussion also festhalten, dass Vertrauen in einer gewissen Form als „persönlichere" Einstellung verstanden werden muss, als es eine sozialkalkulatorische wäre. Was genau unter einer solchen persönlichen Einstellung gemeint sein kann, bleibt damit aber weiterhin rätselhaft. Das Attribut „persönlich" wird hier eher *ad-hoc* und in einem vortheoretischen Sinne verwendet. Es leistet damit noch nichts zur profunden Erläuterung interpersonellen Vertrauens. Dies soll in den folgenden beiden Unterkapiteln geschehen, in denen Vertrauen als eine anerkennende Haltung verständlich gemacht werden soll.

6.2 Anerkennung und Zeugenschaft: Eine Kritik an Richard Moran

In den beiden folgenden Unterkapiteln soll also die These artikuliert und argumentativ begründet werden, dass es sich bei interpersonellen Vertrauen um eine Haltung personaler Anerkennung handelt. Die entscheidenden Punkte werden wir mit Rekurs auf den Diskurs um Zeugenschaft und Sprechervertrauen entwickeln, in dem die zentralen Ideen zur Charakterisierung einer genuin vertrauenden im Gegensatz zu einer bloß kalkulierenden Haltung angelegt sind. Diese Unterscheidung findet im Zeugenschaftsdiskurs seine Analogie in Form des oben angesprochenen Unterschiedes zwischen dem, was es heißt „*zu glauben, was jemand sagt*" und dem, was es heißt „*jemandem zu glauben, was er sagt*" wider. Oben habe ich den Unterschied mithilfe eines kurzen Beispieles illustriert: Wenn ich aufgrund eines evidenzbezogenen Kalküls darauf schließe, dass mein Mitbewohner die Wahrheit sagt, dann glaube ich zwar, was er sagt; ich glaube jedoch nicht *ihm*, was er sagt.

In Bezug aber auf die Frage, was genau es heißt jemandem zu vertrauen, stellt der Aufsatz *Getting Told and Being Believed*[210] von Richard Moran einen vielbeachteten Antwortversuch dar, mit dem wir uns in diesem Unterkapitel ausführlicher auseinandersetzen werden. Wie wir zeigen wollen, weist der Ansatz von Moran zwar in die richtige Richtung, indem er zwar zumindest nominal Vertrauen in den Sprecher als eine Form personaler Anerkennung ausweist. Allerdings nimmt er keine präzise Erläuterung des Anerkennungsbegriffes vor. Wie wir zeigen wollen, ist sein

210Moran 2005.

Ansatz deshalb letztlich an den entscheidenden Stellen wenig aussagekräftig.

Das Herzstück des Ansatzes von Moran bildet seine Unterscheidung zwischen der sogenannten *evidential view* und der *assurance view (of testimony)*, die zwei grundlegend verschiedene Arten von Perspektiven bezeichnen, die auf die Mitteilung einer anderen Person eingenommen werden können. Die *evidential view* entspricht dabei der Perspektive, die ich in dem Beispiel gegenüber meinem Mitbewohner einnehme. Kennzeichnend für diese ist, dass die Mitteilung des Anderen allein als besondere Art eines Beweisstückes für die Wahrheit des Gesagten fungiert – in dem Beispiel etwa stellt für mich die Tatsache, dass mir mein Mitbewohner mitteilt, dass p der Fall ist, lediglich einen *Umstand* dar, der nahelegt, dass p der Fall ist.

In der Zeugenschaftsdebatte wird die *evidential view* naheliegenderweise mit sogenannten reduktionistischen Positionen in Verbindung gebracht. Diese besagen, dass Zeugenschaft, neben etwa Wahrnehmung und Erinnerung, keine eigenständige Erkenntnisquelle sei. Die Rechtfertigung von Wissen durch die Zeugenschaft anderer sei letzten Endes vielmehr auf eine der herkömmlichen Erkenntnisquelle (typischerweise sinnliche Wahrnehmung) zurückzuführen. Als historischer Ausgangspunkt für diese Sichtweise werden klassischerweise David Humes Überlegungen zur epistemischen Validität von Zeugenschaft rekonstruiert. So entwickelt Hume eine erfahrungsbasierte Erklärung dafür, wie es gerechtfertigt sein kann, der Mitteilung einer anderen Person Glauben zu schenken.[211] Er sieht die Rechtfertigungsbasis für Vertrauen in Zeugenschaft letztendlich in einer empirisch feststellbaren Korrelation zwischen dem, was Menschen mitteilen und den tatsächlichen Fakten. Hume erläutert zunächst ein allgemeines epistemisches Prinzip, nach dem man bei einer verlässlich *empirisch bestätigten* Korrelationsbeziehung zwischen zweier Ereignistypen beim Auftreten eines Ereignisses des einen Typs auf ein Ereignis des anderen Typs schließen könne. Das Vorgehen nach diesem Prinzip wendet er dann auf den Fall der Zeugenschaft an:

> „It being a general maxim, that no objects have any discoverable connexion together, and that all the inferences, which we can draw from one to another, are founded merely on our experience of their constant and regular conjunction; it is evident that we ought not to make an exception to this maxim in favour of human testimony, whose connexion with any event seems, in itself, as little necessary as any other.“[212]

Aus der wiederholten Beobachtung, dass das, was Menschen sagen der Wirklichkeit

211 Ein Anlass für Hume darüber nachzudenken ist dabei bei ihm die Frage, ob man Zeugenschaftsberichten über Wunder Glauben schenken dürfe (Hume 2007b, Sect. X). Für eine historische Einordnung dieses Argumentes, siehe: McMyler 2011, Kap.1.

212 Hume 2007b, 81.

entspricht, könne man induktiv schließen, dass ihre Mitteilungen üblicherweise korrekt sind. Dieser induktive Zusammenhang könne natürlich zudem beliebig genau spezifiziert werden: So mag sich herausstellen, dass es bestimmte Personen oder Typen von Personen gibt, die besonders verlässlich die Wahrheit sagen. Womöglich lässt sich auch noch präziser eruieren, dass sie dies unter bestimmten Typen von Kontexten oder in Bezug auf gewisse Gegenstandsbereiche tun. Für den über empirische Erfahrung und Induktion feststellbaren Zusammenhang zwischen Bezeugung und bezeugter Tatsache können also durchaus vielfältig spezifizierte Geltungsbedingungen herausgefunden werden.

Gemäß dieser – zugegebenermaßen für unsere Zwecke vereinfachte[213]– Rekonstruktion der Position Humes werden also konkrete Mitteilungen einer Person als ein Indiz dafür aufgefasst, dass das, was sie sagt, tatsächlich auch der Fall ist. Einer Mitteilung zu glauben stellt entsprechend ein Schließen von einem empirisch gegebenem Faktum auf ein anderes nach Maßgabe empirisch-induktiv gewonnener Regeln dar. Die Berichte anderer sind folgerichtig auch nicht als eigenständige Erkenntnisquelle anzusehen. Vielmehr beruht ihre epistemische Validität allein auf den durch *sinnliche Wahrnehmung* ermittelten Korrelationsbeziehungen, nach denen man im konkreten Fall auf die Wahrheit des Gesagten schließen kann.

Wohlgemerkt ist damit übrigens nicht gesagt, dass es sich bei diesem Schließen um einen ausdrücklichen Denkprozess handeln muss. Im Gegenteil kann dieser gerade unter Bezugnahme auf Humes eigener Induktionstheorie als ein automatisierter Habitualisierungsprozess[214] verstanden werden – ein Thema auf das wir in Kapitel 8 zurückkommen werden.

Die reduktionistische Position Humes stellt allerdings nur eine besonders prominente Version der *evidential view* dar. Für eine andere, und zwar nicht-reduktionistische Fassung steht hingegen C.A.J. Coadys Ansatz.[215] Dieser argumentiert nicht über die *empirische* Bewährtheit von

213 So gibt etwa C.A.J. Coady die Position Humes wieder (Coady 1992, 81f.). Axel Gelfert spricht diesbezüglich von der „«received view» of Hume‘s account of testimony“ (Gelfert 2010, 63). Er hält dieser eine alternative Rekonstruktion entgegen, die auf die Rolle der menschlichen Natur in Hume‘schen Textpassagen aufmerksam macht. In diesen Abschnitten scheint Hume keine reduktionistische Position zu vertreten, sondern eher eine non-reduktionistische, das heißt, er sieht Zeugenschaft *neben* Sinneswahrnehmung als eine eigenständige Erkenntnisquelle an. Siehe hierzu auch den der Standardlesart Humes gegenüber kritischen Beitrag von Paul Faulkner (1998). Dass diese alternative Lesart jedoch im Wesentlichen nichts an Humes erfahrungsbasierten Modell der Rechtfertigung von Vertrauen in Zeugenschaft ändert, zeigt: Kaminski 2020, II, 1.4.

214 Siehe die Erläuterung aus der *Enquiry* dazu, inwiefern Induktion einen Habitualisierungsprozess darstellt: Stellen wir uns eine Person vor, die „immediately infers the existence of one object from the appearance of the other. Yet he has not, by all his experience, acquired any idea or knowledge of the secret power by which the one object produces the other; nor is it, by any process of reasoning, he is engaged to draw this inference. But still he finds himself determined to draw it: And though he should be convinced that his understanding has no part in the operation, he would nevertheless continue in the same course of thinking. There is some other principle which determines him to form such a conclusion. This principle is Custom or Habit“ (Hume 2007b, 31f.). Im Kontext der Zeugenschaftsdiskussion weist auch Miranda Fricker (2007, Fn.2) auf diesen Punkt hin.

215 Coady 1992.

Inferenzregeln für die Verlässlichkeit der Mitteilungen anderer, sondern vielmehr über verschiedene Begründungen, denen er *a priori*-Gültigkeit zuschreibt. Wir können uns an dieser Stelle nicht im Detail mit den verschiedenen Argumenten Coadys auseinandersetzen. In gebündelter Form wiedergegeben, argumentiert Coady, dass wenn es keine weitestgehende Übereinstimmung zwischen den Berichten anderer und den Fakten geben würde, die Praxis des Mitteilens, erstens, keinen praktischen Sinn hätte, zweitens, nicht in sich konsistent gedacht werden könne, und drittens, überhaupt nicht erlernbar wäre.[216] Insofern es jedoch eine Praxis des Mitteilens gibt, könne man davon ausgehen, dass Zeugenschaftsberichte in der Regel wahr seien. Auch diese transzendental zu verstehenden Argumente[217] sind wohlgemerkt der *evidential view* zuzuschlagen. Sie zielen nämlich letztendlich ebenso darauf ab nachzuweisen, dass die Zeugenschaft anderer in der Regel ein verlässliches *Indiz* für die Wahrheit des Gesagten darstellt. Der wesentliche Unterschied zur humeschen Position liegt allein in Bezug auf der Erklärung der *Beweiskraft* dieses Indizes begründet, die eben nicht auf empirischer Beobachtung, sondern transzendentaler Argumentation beruht.

Einer von Morans entscheidenden Einwänden gegen die *evidential view* besteht in der Feststellung, dass mit dieser nicht eingefangen werden kann, was es heißt *jemandem* zu glauben. So kommt in der evidenziellen Perspektive der Sprechakt der anderen Person lediglich als Beweis vor. Dies berücksichtigt jedoch nicht die Spezifität dessen, was es heißt jemandem etwas mitzuteilen bzw. jemandem zu glauben. Während Moran die *evidential view* schwerpunktmäßig auf diesem Wege kritisierte, haben vor allen Dingen im Anschluss an seinen Aufsatz andere Autor:innen auf die irritierend verdinglichte Rolle abgehoben, die in der *evidential view* dem Sprecher zukommt. Sofern die Mitteilung des Sprechers als bloßer Umstand verstanden wird, durch den sich auf die Wahrheit des Gesagten schließen lässt, komme dem Sprecher lediglich die Position eines „truth-gauge“[218] zu. Dies scheint jedoch mit einer Einstellung des *Vertrauens* in die andere Person unvereinbar, wie etwa Elizabeth Fricker scharf auf den Punkt bringt: „We do not regard or treat others as mere natural signs of what they assert, nor merely *as* measuring instruments, when we believe what they tell us through our trust in them.“[219] Dies bringt deutlich zum Ausdruck, dass wir es bei der *evidential view* wiederum nur mit einer besonderen Version einer sozialkalkulierenden Einstellung zu tun haben.

216Vgl. ebd., 85-93.
217Vgl. Kaminski 2020, 75-88.
218Hinchman 2005, 580.
219Fricker 2006, 602. Arnon Keren nennt dies auch das „perversity argument“ gegen sogenannte evidenzielle Positionen in der Zeugenschaftsdebatte (Keren 2012, 700).

Spiegelbildlich machen auch Vertreter:innen einer nicht-reduktionistischen Position in Bezug auf Zeugenschaft zuweilen darauf aufmerksam, dass aus der Sicht der mitteilenden Person, die eigene Aussage nicht als Bereitstellung eines bloßen Beweises gesehen wird und auch nicht so vom Adressierten gesehen werden soll. Es wäre entsprechend ungewöhnlich, wenn ich jemandem mitteilen würde, dass p der Fall ist, um anschließend zu kommentieren: „Jetzt habe ich es dir gesagt, mach mit der Aussage, was du willst!“ [220]

Mit dieser Problematisierung ist jedoch noch nicht positiv beantwortet, was genau es demgegenüber heißt *jemandem* zu glauben. Nach Moran stellt dies eine prinzipiell andere Art von Haltung gegenüber einem Sprecher dar, als sie in der *evidential view* verkörpert ist. Seine zentrale These ist die folgende: Insofern man jemandem glaubt, wird die Zeugenschaft des anderen nicht mehr als bloßer Beweis, sondern als eine *Versicherung* des Sprechers aufgefasst – dies bezeichnet er als *assurance view of testimony*. Um den Unterschied zwischen beiden Perspektiven zu verdeutlichen, hebt Moran insbesondere auf den unterschiedlichen Stellenwert ab, die jeweils der *Absicht* des Sprechers zukommt. In der evidentialistischen Sichtweise wird schließlich die Mitteilung des Anderen analog einem Beweisstück aufgefasst. Ausschlaggebend für den epistemischen Wert von Beweisen sei aber typischerweise, dass diese gerade nicht fingiert, also *absichtsvoll* fabriziert sind: „Ordinarily, if I confront something as evidence (the telltale footprint, the cigartette butt left in the ashtray) and then learn that it was left there deliberately, even with the intention of bringing me to a particuar belief, this will only discredit it as evidence in my eyes.“[221] Das Gegenteil gelte jedoch in der *assurance view*, also für den Fall, in dem wir nicht bloß glauben, dass eine Person die Wahrheit sagt, sondern vielmehr *der Person glauben*, dass sie die Wahrheit sagt. In kritischer Weiterentwicklung von Überlegungen von Paul Grice[222], betont Moran, dass für diese entscheidend ist, die Mitteilung des Sprechers als *absichtliche* Handlungen aufzufassen.

So hielt Grice fest, dass für eine Mitteilung (dass p der Fall ist) nicht nur entscheidend sei, dass der Sprecher die Absicht habe, den Hörer p glauben zu lassen. Diese Absicht habe ich schließlich auch, wenn ich etwa einer anderen Person absichtlich Beweise für p unterbreite. Der Unterschied zwischen der absichtlichen Lieferung von Beweisen und einer Mitteilung[223] ist aber, dass die Absicht im zweiten Fall eine *notwendige* Rolle spielt. Die Beweise haben ihre Überzeugungskraft unabhängig davon, ob ich sie bewusst einer Person vorlege, oder ob diese sie

220 Siehe die Bemerkung von Moran: „as if the meaning of his utterance were «Now I have spoken; make of it what you will» rather than «Take it from me»“ (Moran 2005, 26).
221 Moran 2006, 6.
222 Vgl. Grice 1957.
223 So zieht Grice einen Unterschied zwischen *openly and deliberately letting someone know* und *telling someone* (Grice 1957, 382).

selber findet. Meine Absicht besteht dann lediglich darin, einer anderen Person die Beweise zugänglich zu machen. Sie überzeugen als Beweise aber gerade unabhängig von meiner Überzeugungsabsicht. Bei einer Mitteilung verhält es sich anders: Deren Überzeugungskraft hängt *konstitutiv* damit zusammen, dass sie freiwillig und absichtsvoll vorgenommen wurde. Eine Mitteilung ist nur insofern überzeugend, als dahinter die freie Absicht des Mitteilenden anerkannt wird, die andere Person zu überzeugen.[224]

Von diesem Gedanken ausgehend, fragt sich Moran jedoch weitergehend, wie dies aus der Perspektive des Zuhörers funktionieren kann. Wie kann man durch Anerkennung der bloßen *Absicht* des Sprechers, dem Zuhörer etwas Bestimmtes glauben zu lassen, dazu kommen dies *tatsächlich* zu glauben? Immerhin scheint doch generell zu gelten, dass der Umstand, dass jemand von mir will, dass ich ϕ tue, mir an sich noch kein Anlass dafür gibt, tatsächlich ϕ zu tun. In Grices Überlegung verbleibt also noch eine Leerstelle. Um diese zu füllen, hebt Moran auf den *Akt der Mitteilung* als einem bestimmten öffentlichen Akt ab: „The epistemic value of his words is something publicly conferred on them by the speaker, by presenting his utterance *as an assertion.*"[225] Diese Überlegungen weisen den Sprechakt als eine bestimmte Art dessen aus, was wir oben als öffentliches *commitment* diskutiert haben. Schließlich haben wir ein *commitment* als einen öffentlichen Akt definiert, der *qua* der gegebenen Konventionen und Praxis, als ein Akt der freiwilligen normativen Bindung verstanden wird. Ebenso wie die Unterzeichnung eines Vertrages stelle eine Mitteilung „by its nature"[226] die offene Verantwortungsübernahme des Sprechers für das Gesagte dar.

Erst vor diesem Hintergrund des Verständnisses eines Aktes der Mitteilung lässt sich die Bedeutung der Absicht des Sprechers verstehen. Um eine genuine verpflichtende Mitteilung kann es sich dabei nämlich nur handeln, wenn es sich dabei um einen absichtlich vollzogenen Sprechakt handelt. Ebenso wie man nur insofern etwas versprechen kann, als man die Absicht hat, ein Versprechen abzugeben und sich damit normativ zu binden, kann man auch in einem basalen Sinne *nur absichtlich* etwas mitteilen.[227] Die Absicht ist hierbei insofern entscheidend, als sie konstitutiv für den Akt der Mitteilung selber ist. Etwas unabsichtlich (sagen wir etwa unter Hypnose) mitzuteilen, wäre schließlich keine wirkliche Mitteilung, oder wie Moran schreibt: „But nothing

224Diesem Gedanken folgend hält Grice für die Perspektive eines Sprechers A fest: „A uttered x with the intention of inducing a belief by means of the recognition of this intention." (ebd., 384).

225Moran 2005, 15.

226Ebd., 16.

227„An utterance counts as an assertion or an apology just in case the speaker presents it as such to his audience, in the appropriate context where his audience can be expected to recognize what is being offered. The speaker cannot count as having promised or asserted something if he had no such intention, or if he did not present his utterance to be seen as a promise or assertion." (ebd., 21).

can count as someone's assurance that was not freely presented as such, just as talking in one's sleep cannot count as making an assertion or a promise."[228]

Der letztendliche Clou der Überlegung Morans besteht nun darin hervorzuheben, dass der Sprecher im Zuge des Sprechaktes dem Hörer einen besonderen Grund dafür gibt das Mitgeteilte zu glauben. Diesen Grund sieht er als kategorial von einem bloß evidentiellen Grund zu unterscheiden an. Schließlich hängt dessen Validität, dass der Sprecher durch seine Mitteilung eine Verantwortung übernimmt, er also für die Wahrheit des Gesagten bürgt. Dies ist die entscheidende Art von Grund, die sowohl der Sprecher dem Hörer zu geben beabsichtigt, als auch für den Hörer ausschlaggebend ist, um dem Sprecher zu glauben. Der Sprecher „must present his action as being without epistemic significance apart from his explicit assumption of responsibility for that significance."[229] Der Hörer wiederum akzeptiert die Garantie des Sprechers und nimmt diese als einen Grund für das Mitgeteilte. Dieser Grund qua der öffentlichen Verantwortungsübernahme stellt nach Moran den entscheidenden Kontrast zur evidenziellen Sichtweise auf Mitteilungen dar.[230]

Moran hält also fest, dass die evidenzielle Sichtweise kategorial ungeeignet ist, um auf den Begriff zu bringen, was es heißt jemandem zu vertrauen. Allerdings werde ich im Folgenden Zweifel formulieren, ob die *assurance view* dazu tatsächlich besser geeignet ist. Um die Kritik an Moran einzuleiten, werde ich zunächst eine Dilemma bei Moran aufzeigen, das sich in Bezug auf die Frage ergibt, für was genau der Sprecher eigentlich Verantwortung trägt. An diesem Punkt nämlich ist Moran nicht eindeutig, und schillert zwischen zwei verschiedenen Positionen hin und her: Zuweilen scheint er hervorzuheben, dass der Sprecher für die *Überzeugung* desjenigen, der ihm glaubt bzw. dem Grund den er ihm gibt, verantwortlich sei.[231] An anderer Stelle formuliert er, dass der Sprecher für die *Wahrheit* des Gesagten verantwortlich sei.[232] Das Dilemma zeigt sich darin, dass sich die erste Lesart in einer Zirkularität zu verstricken droht, in der nicht mehr ersichtlich wird, wie die Mitteilung dem Hörer einen Grund geben kann. In der zweiten Lesart jedoch kann Moran nicht deutlich machen, inwiefern die *asurance view* letztlich nicht doch eine besondere Form der evidenziellen Sichtweise darstellt.

Um den ersten Ast des Dilemmas darzulegen, nehmen wir zunächst einmal an, dass Moran davon ausgeht, dass der Sprecher eine Verantwortung für die Überzeugung des Hörers hat. Diese

228Ebd., 7.
229Ebd., 14.
230Siehe auch die Weiterentwicklung des Ansatzes von Benjamin McMyler (2011).
231Moran 2005, 6,17, 18, 22.
232Ebd.,18, 19, 20, 22, 23.

Überzeugung jedoch soll der Hörer, wie wir gesehen haben, durch einen bestimmten Grund gewonnen haben: Nämlich auf Grundlage der Verantwortungsübernahme des Sprechers. Dies bedeutet es nach Moran einem Sprecher zu glauben. Wie kann dies aber geschehen? Der Sprecher ist verantwortlich für eine Überzeugung, die der andere erst aus dem Grund haben soll, dass der Sprecher für die Überzeugung verantwortlich ist. Es scheint sich damit für diese Lesart der Verantwortung des Sprechers eine viziöse Zirkularität zu ergeben: Der Grund, auf dessen *Grundlage* der Sprecher glauben soll, ist einer, der erst nur dann bestehen kann, wenn der Sprecher *bereits* glaubt.[233]

Ein möglicher Ausweg aus diesem Zirkel kann in der Annahme bestehen, dass der Mitteilende primär nicht für die Überzeugung des Adressierten, sondern schlichtweg für die *Wahrheit* des Gesagten trägt. Dies ist nicht nur in der Zeugenschaftsdebatte eine allgemein verbreitete Ansicht[234], Moran selbst scheint diese gar in manchen seiner Formulierungen zu vertreten: „On the Assurance View, dependence on someone's freely assuming responsibility *for the truth of P*, presenting himself as a kind of guarantor, provides me with a characteristic reason to believe, different in kind from anything provided by evidence alone."[235] Dies bedeutet, dass sich der Sprecher mit seiner Mitteilung verantwortlich macht, *unabhängig* davon, ob der Adressierte ihm glaubt oder nicht. Ganz unabhängig vom Zirkularitätsproblem der ersten Lesart scheint dies auch schlichtweg die zutreffendere Beschreibung zu sein: Wenn ich jemanden mitteile, dass p, dann ändere ich damit *allein qua Mitteilungsakt* die normative Beziehung zwischen uns beiden. Dies zeigt sich auch darin, dass der andere mir Vorwürfe machen kann, wenn ich die Unwahrheit sage, und zwar unabhängig davon, ob er mir tatsächlich geglaubt hat oder nicht.[236] [237] Zumindest in erster Instanz mache ich mich dem Adressierten gegenüber allein für die Wahrheit des Gesagten

233 Der Zirkel erinnert an den Zirkel in Scanlons Versuch (1998, 295-327) den verpflichtenden Charakter von Versprechen über die damit im Versprechensnehmer geweckte *Erwartung* zu begründen. Auch hier gilt, dass der verpflichtende Charakter eines Versprechens nicht auf einer Erwartung beruhen kann. Diese Erwartung nämlich kann umgekehrt erst dadurch erklärt werden, dass sich der Versprechensgeber verpflichtet hat. Siehe hierzu auch: Kolodny/Wallace 2003.

234 Siehe beispielhaft die Erläuterung von Timothy Williamson: "To make an assertion is to confer a responsibility (on oneself) for the truth of its content; to satisfy the rule of assertion, by having the requisite knowledge, is to discharge that responsibility, by epistemically ensuring the truth of the content." (Williamson 2000, 268f.). In Auseinandersetzung mit Moran argumentiert ebenso Thomas Simpson für diese Sichtweise (2018, 6-8).

235 Moran 2005, 7; Hervorhebung von mir.

236 Allerdings versucht Moran auch einen Unterschied zwischen dem adressierten Zuhörer und einem bloß Mithörenden zu ziehen, was dafür spricht, dass er vermutlich der ersten Lesart einen Punkt einräumt. Ob es ihm gelingt, diesen Unterschied herauszuarbeiten, kann jedoch hinterfragt werden (siehe auch die erhellenden Überlegungen von Edward Hinchman (2014, 37-39) dazu). Die erste Lesart pflanzt sich im Übrigen, soweit ich sehe in Benajmin McMylers an Moran anknüpfenden Überlegungen fort (vgl. McMyler 2011, 102).

237 In erster Instanz macht sich der Sprecher also für die Wahrheit des Gesagten verantwortlich. Das schließt nicht aus, dass er darüber hinaus *besondere* Verantwortlichkeiten *gegenüber* dem Hörer hat. Ein interessanter Ausgangspunkt, um über diese Fragen nachzudenken bietet: Hinchman 2014.

verantwortlich, nicht für das, was er daraufhin glaubt. Hier sehen wir auch, dass sich eine Sichtweise bestätigt, für die wir oben argumentiert haben, nämlich die, dass die normative Erwartung im Vertrauen nur begleitenden Charakter hat.

Wenn wir nun aber diese zweite Lesart der Verantwortungsübernahme des Sprechenden voraussetzen, drängt sich die tieferliegende Frage auf, was es bedeutet diesen Akt der Verantwortungsübernahme nicht *selber* wieder als einen evidentiellen Grund aufzufassen. Moran selbst wirft diese Frage am Ende seines Aufsatzes auf: Könnte es schließlich nicht sein, dass der *Umstand*, *dass* der Mitteilende versichert, dass das Mitgeteilte wahr ist, und er sich entsprechend verantwortlich macht, letztlich nur als eine *besondere Form von Evidenz* aufgefasst werden kann, die dafür spricht, dass das Mitgeteilte tatsächlich der Fall ist? So mag man sich etwa vorstellen, dass ein Hörer folgende Überlegung anstellt: „Sprecher S versichert mir, dass p der Fall ist. Damit übernimmt er die Verantwortung dafür, dass p wahr ist. Will er dieser Verantwortung gerecht werden, hat er somit einen speziellen Grund die Wahrheit zu sagen. Und daraus *schließe* ich, dass p tatsächlich der Fall ist“. Diese Überlegung setzt zwar voraus, dass man die Äußerung des Sprechers als eine Versicherung erkennt[238], läuft aber letztendlich doch bloß auf eine *besondere Art* evidenziellen Schließens hinaus. Evidenzielles Schließen charakterisiert Moran immerhin als „just a matter of an inference from behavior which is seen as rational to some conclusion about the state of the world.“[239] Eben einen solchen Schluss scheint der Hörer in dem Beispielfall zu vollziehen. Die Gültigkeit des Schlusses beruht zwar darauf, dass man es mit einer Versicherung des Sprechers zu tun hat. Diese fungiert hier jedoch letztendlich nur als besonderes Element im Rahmen eines evidenziellen Kalküls.

Es ist wichtig zu sehen, dass dieses Problem auf das genau Verständnis der Haltung des Hörers abzielt. Oben haben wir mithilfe von Moran recht ausführlich die Perspektive des Sprechers beschrieben, und welche Bedeutung in diesem Zusammenhang dem Sprechakt der Mitteilung zukommt. Nun geht es darum genauer die Einstellung eines Hörers nachzuvollziehen. Was heißt es, wenn ein Hörer *einem Sprecher* glaubt?

Auf diesen Einwand liefert Moran eine ausführliche Entgegnung, die jedoch offenbart, welche Erklärungslücken sein Ansatz der *assurance view* aufweist. So postuliert Moran gegen den

238Dies ist die gleiche Perspektive, die auch ein Mithörer gegenüber dem Gesagten einnehmen könnte (siehe den Vorwurf von Lackey 2008, 233). In dieser evidenziellen Perspektive ließe sich auch die Beweiskraft mehrerer Mitteilungen anderer schlichtweg addieren. Auf diesen Umstand macht etwa Stanley Cavell aufmerksam: „[…] suppose I believe what is said to me only if two say it. A appears and says «X is Y». I think: That's what *you* say. Then B appears and confirms it, i.e., repeats it. Now I believe, or at any rate am prepared to say, that X is Y. But I do not believe either A or B, any more than I believe my weathervane when I read the wind's direction from it.“ (Cavell 1979, 391).

239Moran 2005, 4.

Vorbehalt gerichtet zunächst eine sogenannte Korrespondenznorm (*norm of correspondence*) zwischen dem Grund, den der Sprecher dem Hörer zu geben beabsichtigt und dem Grund aus dem heraus der Hörer dem Sprecher glaubt. Fassen wir dies in einem etwas längeren Zitat diesen Gedanken zusammen:

> „When an act of telling completes itself, speaker and audience are aligned [...] through their mutual recognition of the speaker's role in determining the kind of reason for belief that is up for acceptance, so that when the speaker is believed there is a non-accidental relation between the reason presented and the reason accepted. The speaker says, in effect, «The kind of reason for belief you gain from my statement is precisely the kind of reason for belief I am hereby presenting myself as offering you. [...]»".[240]

Es bleiben aber Zweifel, ob es Moran damit gelingt ein nicht-evidenzielles Verständnis dessen, was es heißt jemandem zu glauben *deutlich zu machen*. So markiert die Feststellung, dass der vom Sprecher präsentierte und vom Hörer akzeptierte Grund einen Grund gleicher Art darstellen, zunächst einmal nur ein formales Merkmal: Es muss für beide der *gleiche* Grund sein. Die substanzielle Frage bleibt jedoch bestehen, wie genau dieser Grund und die Perspektive des Hörers damit genau zu denken sind. Es ist der gleiche Grund auf beiden Seiten, aber um was für eine Art von Grund handelt es sich dabei? Die Korrespondenznorm markiert in diesem Sinne höchstens eine notwendige formale Bedingung, die gegeben sein muss, damit man im eigentlichen Sinne davon sprechen könne, dass jemand jemandem etwas mitgeteilt hat. Sie ist in der Formulierung Morans aber nicht hinreichend und erläutert damit nicht substanziell, was es heißt jemandem etwa mitzuteilen bzw. jemanden zu vertrauen.

Zwar unterstreicht Moran, *dass* wir weder aus der Perspektive des Mitteilenden noch aus der Perspektive desjenigen, der dem Sprecher glaubt, die Mitteilung als Bereitstellung eines zu berücksichtigenden Beweises verstehen. Auf diesen Punkt haben wir bereits oben hingewiesen: Es wäre merkwürdig, wenn mir jemand mitteilt, dass p der Fall ist, und anschließend hinterherschiebt: „Jetzt mache mit dieser Aussage, was du willst". Etwas mitzuteilen zielt auf etwas grundlegend anderes ab, als das Bereitstellen eines prinzipiell anfechtbaren Beweises. Die entscheidende Frage aber bleibt: auf was *genau* zielt sie dann *positiv* gesehen ab? Und was heißt dementsprechend aus Sicht des Hörers, dem Sprecher zu vertrauen?

An anderen Stellen wiederholt Moran lediglich den Punkt, dass in der *assurance view* der Hörer einen Grund hat, der mit der Verantwortungsübernahme des Sprechers zusammenhängt: Der

240Moran 2005, 25.

Grund sei „bound up with the speaker presenting himself as accountable for the truth.“[241] Erinnern wir uns aber an das Problem von oben: Die Verantwortungsübernahme des Sprechers besteht erst einmal unabhängig davon, ob und in welcher Form der Hörer dem Sprecher auch glaubt. Mehr noch, ein Hörer kann sich gerade in rein evidenzieller Manier auf die Verantwortungsübernahme des Sprechers beziehen, wie wir oben dargelegt haben. Ohne weitere Erläuterung bleibt der Hinweis auf die Verantwortlichkeit des Sprechers also wenig hilfreich um gehaltvoll auf den Begriff zu bringen, was es heißt jemandem zu glauben.

Diesen Überlegungen folgend bleibt bei Morans Vorschlag also weiterhin dunkel in Bezug auf die Frage, was es eigentlich heißt einem Sprecher zu vertrauen. Diesbezüglich benennt Moran lediglich ein notwendiges formales Merkmal in Form der Korrespondenznorm. Im Folgenden werden wir den Versuch unternehmen eine klarere Charakterisierung dessen was es heißt jemandem zu vertrauen zu geben. Wir werden damit der Vermutung Morans nachspüren, dass einem Sprecher zu glauben bedeutet, ihm gegenüber eine besondere Form von Haltung, eine *anerkennende* Haltung einzunehmen. Allerdings werden wir diese Idee anders ausarbeiten, auch mit einer profunden Klärung dessen, was es überhaupt heißt jemanden *anzuerkennen* – eine Klärung, die bei Moran ausbleibt, der sich eher auf einen vortheoretischen Anerkennungsbegriff bezieht.

6.3 Vertrauen als Anerkennung

Die Diskussion des Ansatzes von Moran führte vor Augen, dass es sich auch bei einer scharfen Abgrenzung von der *evidential view*, damit noch nicht von selbst versteht, was es positiv ausbuchstabiert heißt, einem Sprecher zu glauben. Der Vorschlag den wir hier entwickeln werden lautet, dass mit der Mitteilung des Sprechers ein impliziter Gesuch nach *Anerkennung als vertrauenswürdiger Sprecher* einhergeht. Umgekehrt bedeutet es einem Sprecher zu glauben, dieses Gesuch zu akzeptieren. Wir werden dies die *Anerkennungsthese des Vertrauens* nennen. Wohlgemerkt zielt demnach der Sprecher nicht bloß darauf ab, Verantwortung für das Gesagte zu übernehmen – dies geschieht im Grunde ohnehin, unabhängig davon, ob der Hörer ihm letztlich glaubt oder nicht. Er zielt im Normalfall[242] darauf ab, dass der Hörer ihn *ernst nimmt* und ihm

241Ebd., 25.

242Wir gehen hier vom mit Austin gesprochen gelungenen (*felicitious*) Mitteilungsakt aus (Vgl. Fricker 2006, 596). Ein Sprecher kann zwar immer auch eine bloß täuschende Absicht haben. Im Normalfall aber, der uns hier interessiert, geht er davon aus, was er sagt zu wissen und beabsichtigt dieses Wissen mitzuteilen.

aufgrund dessen glaubt. Der Hörer wiederum muss sich irreduzibel in der einen oder anderen Form zu diesem Gesuch verhalten. Dem Sprecher zu vertrauen bedeutet dabei auf dieses implizite Gesuch einzugehen und ihn als vertrauenswürdig anzuerkennen. Einer Person zu glauben bedeutet entsprechend sie ernst zu nehmen. Wie Elizabeth Fricker an einer Stelle festhält, ist es eine Form von Respekt.[243] Erst mithilfe dieser Idee, so die hier entworfene These, kann angemessen verstanden werden, was es heißt einem Sprecher zu vertrauen.

6.3.1 Die Anerkennungsthese des Vertrauens

Zur schrittweisen Erläuterung der Anerkennungsthese gehen wir zunächst auf das implizite Anerkennungsgesuch des Sprechers ein. Moran arbeitet deutlich heraus, dass der Sprecher mit seinem Sprechakt nicht darauf abzielt, dass dieser als bloßes Indiz für das Gesagte aufgefasst wird. Wenn ich jemandem sage „Ich gebe dir mein Wort: p ist der Fall“, so ist der unterliegende Appell dieser Aussage nicht: „Und jetzt mach mit dieser Aussage, was du willst!“. Im Einklang mit diesem Gedanken kann man als Sprecher auch empört reagieren, wenn der Zuhörer das, was man ihm mitteilt so behandelt.[244] Diesen Punkt versuchte er auch mit der erwähnten Korrespondenznorm zu berücksichtigen.

Die entscheidende Frage, die damit aber bleibt, lautet, *was genau* der Sprecher dann mit der Mitteilung tatsächlich *beabsichtigt*. Wir werden hier den Vorschlag machen und weiter unten entwickleln, dass dem Sprechakt ein impliziter Appell der Art „Nimm mich Ernst!“ unterliegt.[245] Dieser, so der hier eingeführte Vorschlag, sollte präziser verstanden werden als Ausdruck eines noch näher zu bestimmenden *Anerkennungsgesuches* an den Adressierten.

Dass dem so ist, deutet sich symptomatisch besonders klar in der Signifikanz verweigerten

243 Fricker führt diesen Punkt leider nicht viel weitergehend aus: “When I take another’s word for it that *p* […] I treat my teller with respect, in a way that I do not when I treat her expressed belief merely as defeasible evidence. One might say that I treat her as an end, not merely as a means” (Fricker 2006, 607).

244 Siehe die Feststellung von Edward Hinchman: „It’s an abuse of [a teller‘s] status as addressee if [a herarer] treats [his] telling as one he has merely heard or overheard, that is, if he treats it as a mere assertion.“ (Hinchman 2005 , 566).

245 Eine interessante Beobachtung von Arnon Keren unterstützt diese Idee: „Consider […] a speaker, Sandy, who addresses us and says: “take my word for it, *p*.” And suppose that we form the belief that *p*, and that we do so without weighing other relevant evidence available to us. It seems that Sandy cannot now turn to us and say: “You are correct. But you should have considered all the evidence available to you before reaching that conclusion.” Criticizing us in this way appears incompatible with her expectation of us that we take her word for it. If she were to criticize us in this way, she would seem to be withdrawing her original appeal to us that we take her word for it.“ (Keren 2014, 2600).

Vertrauens an. Verweigertes Vertrauen kann schließlich als mangelnde Anerkennung empfunden werden, die eine entsprechende emotionale Komponente aufweist und auch auf die Qualität der Beziehung zur anderen Person übergreift. Auf diese Bedeutung des möglichen Versagens von Vertrauen weist Oli Lagerspetz in einer Textstelle eindrücklich hin:

> „The mere fact of being addressed by another means that he is exposed to us; it is in our power either to accept him or to reject him. We may ignore him or ridicule him. We may be obtuse or question his motives. In all these cases, we are of course reacting to a certain string of words. Yet they involve a rejection not just of the words but of the person who utters them.“[246]

Hier deutet sich an, dass eine Mitteilung mit einem impliziten Anerkennungsgesuch einhergeht, das sich insbesondere in den möglichen Auswirkungen der Zurückweisung dieses Gesuches zeigt. Die Zurückweisung ist eine bestimmte Art der adressierten Person, sich gegenüber dem Sprecher zu verhalten und hat damit Implikationen auf ihre Beziehung. Auch G.E.M. Anscombe weist bereits in ihrem frühen Aufsatz *What is it to Believe Someone?* auf diesen Umstand hin. Sie hält fest: „It is an insult and it may be an injury not to be believed.“[247] [248] Auch diese Beobachtung bestätigt, dass sich mangelndes Vertrauen in einen Sprecher in entsprechenden reaktiven Emotionen zeigt, die als Ausdruck einer Erfahrung fehlender Anerkennung verstanden werden können.

Diese Überlegungen geben uns ein erstes Indiz dafür, dass der Mitteilungsakt eines Sprechers auch von einer Anerkennungsdimension unterlegt ist und das Vertrauen eine entsprechende Reaktion darauf darstellt. Die daran anknüpfenden Anschlussfragen, die wir im Folgenden angehen werden, lauten nun jedoch:

1) Was heißt Anerkennung prinzipiell im hier einschlägigen Sinne?
2) In welcher Hinsicht genau wird eine Person anerkannt, sofern man dieser vertraut?

Gehen wir zunächst auf die erste Frage ein und erläutern, wie in diesem Zusammenhang der Begriff der Anerkennung verstanden werden soll. Dabei ist zunächst einmal darauf hinzuweisen, dass der Anerkennungsbegriff hier im Sinne der *personalen* Anerkennung aufzufassen ist. Einem

246Lagerspetz 1998, 149f.
247Anscombe 2008, 9.
248Interessanterweise zitiert auch Moran diese Stelle am Ende seines Aufsatzes. Der Zusammenhang mit seinen Ausführungen zur *assurance view* scheint er mir aber unanalysiert zu lassen.

weiten Begriffsverständnis nach (ähnlich wie beim Vertrauensbegriff) können nämlich auch Sachverhalte anerkannt werden, etwa wenn ich anerkenne, dass es draußen zu kalt ist, um ohne Mantel aus dem Haus zu gehen, o.ä. Diesen generischen Anerkennungsbegriff haben wir hier jedoch nicht im Sinn, sondern einen, der sich spezifisch auf andere Personen und deren Status bezieht. Die hier relevante Art von Anerkennung besteht grundlegend darin *jemanden* anzuerkennen.

Zugegeben ist dieser personale Anerkennungsbegriff in der philosophischen Diskussion notorisch umstritten, und wird im Einzelnen relativ unterschiedlich aufgefasst. Entsprechend gibt es auch im Diskurs zum Anerkennungsbegriff aktuelle Tendenzen vielmehr von einer Pluralität oder „Familie" von Anerkennungsbegriffen auszugehen, zuweilen verbunden mit dem Versuch deren Binnensystematik herauszuarbeiten.[249] Was über die verschiedenen Konzeptionen hinweg jedoch als ganz allgemeines Merkmal von personaler Anerkennung festgehalten werden kann ist das Folgende: Im grundlegendsten Sinne kann personale Anerkennung verstanden werden als das Einsehen eines Rechtes, eines normativen Anspruches oder einer normativen Erwartung einer Person. Eine Person oder Personengruppe anzuerkennen bedeutet formal erläutert, ihr einen bestimmten normativen Status zuzuerkennen. Dies ist der entscheidende Aspekt, den wir auch für den hier relevanten Anerkennungsbegriff zur Beschreibung einer vertrauenden Haltung hervorheben wollen. Daraus ergibt sich zunächst folgende These, die ich als formale Anerkennungsthese bezeichne:

FORMALE ANERKENNUNGSTHESE

Einer Person zu vertrauen bedeutet, sie anzuerkennen, also einen normativen Anspruch ihrerseits einzusehen.

Dass wir es übrigens tatsächlich mit einer handfesten normativen Erwartung zu tun haben, die vom Sprecher ausgeht, zeigt sich dort besonders deutlich, wo man sich auch bei einem Hörer *beschweren* kann, wenn dieser einem nicht glaubt. Auch diesen Punkt deutet Anscombe an.[250] Dieser Beschwerde nämlich unterliegt eine normative Erwartung an die andere Person, die als

249Vgl. Ikäheimo/Laitinen 2007. Paul Ricouer unterscheidet dabei tatsächlich zwischen 23 verschiedenen Anerkennungsbegriffen (Ricouer 2004, 5–16).

250Anscombe 2008, 9f.

verletzt angesehen wird. Damit wird insgesamt offenbar, dass die normative Situation zwischen den beiden Vertrauenspartnern gegenüber unserer obigen Diskussion komplexer aufzufassen ist. Oben haben wir uns mit der normativen Erwartung des Vertrauenden gegenüber dem Vertrauten auseinandergesetzt. Hier wird deutlich, dass auch eine normative Erwartung des Vertrauensnehmers gegenüber dem Vertrauensgeber besteht.[251]

Die zweite Frage, die wir oben festgehalten habe, zielte darauf ab, genauer den *Inhalt* dieser normativen Erwartung zu verstehen. Was *genau* erwartet der Sprecher vom Hörer? Daraus ergibt sich auch die Antwort auf die Frage, in welcher *Hinsicht* genau der Hörer den Sprecher anerkennt, insofern er diesem vertraut. Dieser Frage können wir uns zunächst wiederum *ex negativo* nähern, indem wir ausschließen, was genau es *nicht* heißt den Sprecher anzuerkennen.

Wiederholt haben wir bereits gesehen, dass der Sprecher nicht einfach nur darauf abzielt, dass man glaubt, was er sagt. Bestimmte Formen zu glauben, dass es der Wahrheit entspricht, was eine Person sagt, stellen gerade keine Form dar, die andere Person als Sprecher ernst zu nehmen. Wie wir gesehen haben, kann man auch das glauben, was eine Person sagt, ohne wirklich der Person zu glauben, was entsprechend als persönliche Verletzung oder Beleidigung aufgefasst werden kann. Stellen wir uns etwa vor der chauvinistische Wissenschaftler X hat die Tendenz die Aussagen und Einschätzungen seiner weiblichen Kolleginnen nicht Ernst zu nehmen. Bei einer Tagung war der von ihm geschätzte männliche Kollege Y verhindert und schickt in Vertretung seine Mitarbeiterin Z, die stellvertretend für Y einen eigenen Vortrag hält. X ist von der Qualität des Vortrages derart beeindruckt, dass er davon ausgeht, es handele sich dabei eigentlich um den ursprünglichen Vortrag von Y, und dass Z diesen bloß in Vertretung vorliest. In diesem Fall bringt er den Aussagen und Einschätzungen des Vortrages großes Vertrauen entgegen, aber: Er vertraut in einer nicht respektvollen Art und Weise nicht *der Vortragenden*. Wenn wir uns vorstellen, dass Z davon erfährt, dass X davon überzeugt ist, dass es sich nicht um ihren eigenen Vortrag handeln kann, mag sie sich beleidigt oder verletzt, in jedem Fall nicht wirklich ernst genommen fühlen.

Anhand des Beispiels lässt sich bereits eine erste Vermutung darüber gewinnen, was es heißt eine Sprecherin ernst zu nehmen. Was der Wissenschaftler X nämlich verweigert, ist Z eine gewisse *Vertrauenswürdigkeit* zuzuschreiben. Die Sprecherin Z nimmt hier für sich in Anspruch als in der Sache vertrauenswürdig aufzutreten. Insofern X dies nicht anerkennt, verweigert er ihr diesen Anspruch. Daraus ergibt sich die spezifische Anerkennungsthese:

251 Für die sozialpsychologische These, dass eine allgemeine Norm zur Anerkennung und Wertschätzung eine Erklärung für das empirisch messbare hohe Maß an Kooperationsbereitschaft (gerade in spieltheoretischen, anonymen Kontexten) liefern können, siehe: Dunning et.al. 2014.

SPEZIFISCHE ANERKENNUNGSTHESE

Einer Person zu vertrauen bedeutet, sie als vertrauenswürdig anzuerkennen, das heißt, den normativen Anspruch ihrerseits einzusehen, als vertrauenswürdig angesehen zu werden.

Der hier unterbreitete Vorschlag lautet also, dass seitens des Vertrauenspartners eine normative Erwartung besteht, ernst genommen, das heißt, als vertrauenswürdig angesehen zu werden. Vertrauen wiederum bedeutet, die andere Person diesbezüglich anzuerkennen, sprich, diese normative Erwartung unmittelbar zu akzeptieren. Die adverbiale Qualifizierung „unmittelbar" soll dabei darauf hinweisen, dass dies nicht wiederum aus einem instrumentellen Hintergedanken heraus geschehen kann, sondern vielmehr aus innerer Einsicht in die Geltung der normativen Anforderung: Eine Person etwa ernst nehmen zu wollen, da man sich dadurch Vorteile verspricht, ist nämlich begrifflich gar nicht möglich. Dies hieße nämlich gerade, sie *nicht wirklich* ernst zu nehmen, sondern nur so zu tun, als ob man sie ernst nehme. Insofern verstehe ich den Terminus der Anerkennung bzw. des Ernstnehmens so, dass er begrifflich gesehen damit einhergeht, dass normative Ansprüche einer Person intrinsisch oder unmittelbar als für das eigene Handeln leitend angeeignet werden.

Zu der normativen Erwartung nach Anerkennung als vertrauenswürdig ist zudem noch zu sagen, dass deren Einlösung nicht zu erzwingen ist. Oben sprachen wir bereits mit Blick auf den Vertrauenden davon, dass die Erfüllung seiner normativen Erwartung nicht einfach, wie ein juristisch begründeter Anspruch, *eingefordert* werden kann. Die Pflichten anderer Personen können eingefordert werden, wenn darunter verstanden wird, dass etwa unter der Androhung von Sanktionen eine Person dazu gebracht wird, das zu tun, was die Pflicht verlangt. Was jedoch nicht unmittelbar forciert werden kann, ist, dass die andere Person sich *vertrauenswürdig* verhält. Unter Vertrauenswürdigkeit haben wir nämlich eine nicht direkt erzwingbare *Haltung* der anderen Person verstanden, die darin besteht die normative Erwartung des Vertrauenden als solche unmittelbar als praktisch wirksamen normativen Maßstab des eigenen Handelns anzuerkennen. In ähnlicher Weise kann hier die normative Erwartung des Vertrauenspartners ebenfalls nicht mit einem erzwingbaren Anspruch gleichgesetzt werden. Dies scheint unmittelbar einzuleuchten: Man kann eine Person immerhin nicht – zumindest nicht direkt – dazu zwingen, jemanden ernst zu nehmen. Dies ist geradezu ein widersprüchliches Unterfangen, denn eine Person allein aus Angst vor Sanktionen

„ernst zu nehmen“, stellt gerade eine Form dar, sie *nicht wirklich* ernst zu nehmen.[252]

6.3.2 Die evaluative Komponente vom Vertrauen

Ausgehend von der Anerkennungstheorie des Vertrauens lässt sich auch deutlich machen, dass Vertrauen eine evaluative Komponente beinhaltet. Unterstützt durch Überlegungen von Peter Strawson und Bennet Helm lässt sich zudem zeigen, dass der Vertrauensgeber und der Vertrauensnehmer auf einer grundlegenden Ebene ein gemeinsames Werturteil teilen. Daraus wiederum ergibt sich auch, dass Vertrauen mit einer Wertschätzung der Vertrauenswürdigkeit des anderen einhergeht. Die begrifflichen Zusammenhänge dazu sollen hier schrittweise kurz skizziert werden.

Vertrauen selber, so haben wir gezeigt, bedeutet einer normativen Anforderung nach Anerkennung oder Respekt der anderen Person nachzukommen. Damit stellt Vertrauen grundlegend betrachtet eine Form von Verhalten aus Einsicht in normative Geltung einer solchen Anforderung dar. Unter Rückgriff auf Überlegungen von Peter Strawson und Bennet Helm lässt sich nun herausstellen, dass dies eine evaluative Komponente enthält, die zudem im Kontext eines weitergehenden komplexen Musters von Evaluationen steht.[253] Helm entwickelt die Zusammenhänge zunächst primär mit Blick auf unsere reaktiven Emotionen und die vielfältigen Weisen wie diese miteinander rational verknüpft sind. Wer etwa empört (*indignant*) ist, wenn einem ein Recht verletzt wird, der erwartet im Normalfall seitens desjenigen, der dieses Recht gebrochen hat Schuldgefühle und seitens unbeteiligter Dritte, dass diese ebenso entrüstet (*resentful*) über die Rechtsverletzung sind. Die verschiedenen reaktiven Emotionen (Empörung, Schuld, Entrüstung) stehen also in einer, in diesem Fall, interpersonellen inferentiellen Beziehung zueinander. Ebenso stellt er aber auch intrapersonelle Beziehungen ähnlicher Art heraus. Wer, um nur einen beispielhaften Zusammenhang herauszugreifen, empört ist, wenn andere ein Recht von ihm missachten, den es jedoch kalt lässt, wenn er dieselben Recht bei anderen verletzte, der verhält sich nicht stringent.[254] Reaktive Einstellungen sind insofern in einem Netzwerk inferentieller Verknüpfungen (*rational pattern*) eingebettet. Sie sind, wie Helm mit einer Formel von Strawson

252Es handelte sich dabei um *Gründe falscher Art* für Anerkennung – ein Terminus, den wir unten genauer erläutern werden.
253Vgl. Helm 2011;2014.
254Ebd., 193f.

festhält, „humanly connected“.[255] Wer von diesen abweicht, der verhält sich, wie Strawson festhält „abnormal“, etwa wie der „moralische Solipsist“, der in egozentrischer Manier nur die eigenen Rechten und Pflichten beachtet, aber ignoriert, ob andere diesen gemäß handeln.

Die reaktiven Emotionen sind aber zugleich durch das bedingt, was uns wichtig ist, was wir valuieren. Wir sind grundlegend gesprochen deswegen empört über das Fehlverhalten anderer, oder haben Schuldgefühle aufgrund unseres eigenen Fehlverhaltens, da wir es als *wichtig* ansehen, dass man sich an normative Anforderungen hält. Gerade deswegen sind wir ja auch vom Fehlverhalten anderer oder uns selbst *emotional* affiziert. Wir stellen Fehlverhalten nicht einfach nur nüchtern fest, sondern sind darüber entrüstet, da es für uns von Bedeutung ist, es uns etwas angeht, dass durch Einsicht in die Geltung von Normen gehandelt wird. Die inferentiellen Zusammenhänge bezüglich der reaktiven Emotionen sind deshalb gleichursprünglich mit inferentiellen Zusammenhängen bezüglich unserer *evaluativen* Urteile in Bezug auf Normen und normativ geleiteten Verhaltens. Die mangelnde Konsistenz etwa des moralischen Solipsisten, der nur über Fehlverhalten ihm gegenüber empört ist, spiegelt damit eine mangelnde Konsistenz bezüglich dessen wider, was ihm wichtig ist, welche Valuierungen er vornimmt. Wer sich in seinem Handeln an dem orientiert, was normativ geboten ist, da er ein solches Verhalten als gut und wichtig ansieht, der sollte etwa konsequenterweise dieses Verhaltens auch *bei anderen* als gut und wichtig ansehen. Daraus folgt beispielsweise ebenso, dass er Verhalten, das gegen normative Anforderungen verstößt als schlecht beurteilt.

Die Implikationen, die sich aus Helms Analyse ergeben, sind dabei noch weiterreichender: So impliziere die positive Bedeutungszuschreibung normativ geleiteten Handelns auch, dass man den normativen Status anderer Personen als wichtig ansieht, die solche Normen für sich einklagen, ebenso, dass man es als relevant ansieht, dass Personen verantwortlich gemacht werden, wenn sie gegen normative Anforderungen verstoßen, dass man es als schlecht ansieht, wenn Verstöße gegen Normen ignoriert werden etc. Die genauen Details der Helm‘schen Analyse werden wir an dieser Stelle beiseite lassen.[256] Stattdessen werden wir nur einen solchen evaluativen inferentiellen Zusammenhang herausgreifen der in Bezug auf das Verhältnis von Vertrauensgeber und -nehmer interessant ist.

Der Vertrauensgeber lässt sich wie gesagt in seinem Verhalten normativ leiten. Insofern man in seinem Verhalten einer normativen Anforderung aber nicht aus instrumentellen Gründen,

255Helm 2014, 192.

256Helm argumentiert darüber hinaus auch für die These, dass das zentrale, wichtige Gut, um das es uns als moralisch eingestellten Akteuren geht, die Gemeinschaft der Akteure selber ist. Aus diesem Gut leiten sich die weiteren Valuationen – die Wichtigkeit der Normen, der normative Status anderer – ab (Ebd).

sondern aus innerer Einsicht in deren Geltung nachkommt, bedeutet dies, dass man es als von generell *positiver Bedeutung* ansieht normativen Anforderungen zu folgen. Intrinsisch den normativen Anforderung nachzukommen, bedeutet trivialerweise, dass man es als *gut* ansieht, sich in seinem Verhalten normativ leiten zu lassen. Hier zeigt sich, dass Vertrauen eine basale evaluative Komponente impliziert. Oben aber haben wir bei der Analyse von Vertrauenswürdigkeit das Gleiche herausgestellt: Auch der Vertrauenswürdige sieht ein Gut darin, das zu tun, was normativ gefordert ist. Beide Akteure teilen deshalb auf einer grundlegenden Ebene beide ein gemeinsames Gut: Ihnen ist beiden am normativ geleiteten Handeln gelegen. Dieser Gedanke mag einen interessanten Ausgangspunkt darstellen, um die häufig geäußerte Grundintuition zu erhellen, dass Vertrauen wesentlich in reziproker Form auftaucht: So suggerieren einige Autor:innen, dass Vertrauen seinen paradigmatischen Ort in Beziehungen *wechselseitigen* Vertrauens habe.[257] Der Punkt sei an dieser Stelle aber nur angedeutet und kann hier nicht weiter verfolgt werden.

Die Helm‘sche Analyse hatte nun zudem ergeben, dass Akteure, die sich in ihrem Verhalten durch normative Einsicht leiten lassen, dieses Verhalten, abgesehen von den „abnormalen“ Fällen, auch *bei anderen* Akteuren hochhalten. Eine Folgerung dieses Zusammenhanges ist, dass der Vertrauensgeber ein Gut in der *Vertrauenswürdigkeit* der anderen Person sieht, da diese eben eine Disposition zum Handeln aus normativer Einsicht darstellt. Wenn Vertrauen als Anerkennung also nur als Verhalten aus normativer Einsicht zu verstehen ist, dann impliziert dies eine Wertschätzung normativ geleiteten Verhaltens generell. Dies wiederum impliziert eine Wertschätzung der Vertrauenswürdigkeit des anderen. Grundlegend zeigt sich also, dass sich die Anerkennung der anderen Person als vertrauenswürdig nicht denken lässt, ohne dem Vertrauenden zu unterstellen, dass er auch einen Wert, eine evaluativ positive Bedeutsamkeit in der Vertrauenswürdigkeit des anderen sieht. Erneut zeigt sich damit: Einer Person zu vertrauen läuft nicht bloß darauf hinaus, eine bloß kognitive Einsicht in deren Vertrauenswürdigkeit zu haben. Vertrauen impliziert, so zeigt sich hier, *auch* eine Valuierung oder Wertschätzung der Vertrauenswürdigkeit des anderen.

Natürlich enthält eine solche Wertschätzung auch eine kognitive Komponente. Samuel Scheffler hat in einer ausführlichen Analyse des Wertschätzungebegriffes herausgearbeitet, dass dieser auch eine Überzeugung impliziert. Wenn ich etwas wertschätze, bedeutet dies schließlich auch, dass ich die Überzeugung habe, dass dieses wertschätzungswürdig ist. Darüber hinaus bedeutet etwas zu wertschätzen jedoch auch, diesem einen Wert zuzuschreiben: „the things we value are things that matter to us, or that we prize or cherish, or that have a certain importance for

257So hält etwa Stephen Darwall fest: „indeed, that I invite you to trust that I am indeed trusting you, to trust in my trust and in me, trusting you“ (Darwall 2017, 42).

us.“[258] Vertrauen in eine Person besteht demnach wesentlich auch in dieser evaluativen Komponente bezüglich der Vertrauenswürdigkeit der anderen Person.

6.3.3 Die Abgrenzung zur sozialkalkulierenden Einstellung

Vor dem Hintergrund der erläuterten These können wir nun auch deutlicher den Unterschied zwischen einer vertrauenden und einer bloß sozialkalkulierenden Einstellung ziehen. Dieser Unterschied ist üblicherweise vor allen Dingen deshalb schwer zu fassen, da beide zunächst einmal darin übereinstimmen, dass mit ihnen der Vertrauenspartner *als vertrauenswürdig angesehen wird.* In dieser Annahme weichen sie nicht voneinander ab. Insofern sie also im Inhalt ihrer bloßen Überzeugung über den Vertrauenspartner übereinstimmen, muss sich der Unterschied zwischen vertrauender und sozialkalkulatorischer Perspektive an einer anderen Stelle zeigen. Er zeigt sich vielmehr am ganzen kategorialen Rahmen, in dem man sich auf die Vertrauenswürdigkeit der anderen Person bezieht. Der Vertrauenswürdigkeit der anderen Person kommt dabei in den jeweiligen Perspektiven ein qualitativ unterschiedlicher Stellenwert zu.

Aus der sozialkalkulierenden Perspektive steht die Realisierung eines Zweckes im Vordergrund, für die sich die Kooperation mit einer anderen Person als probates Mittel anbietet.[259] Im Lichte dieses Interesses ist für den Akteur aber von Belange, dass das gewählte Mittel auch verlässlich ist, was unter anderem durch die Vertrauenswürdigkeit der anderen Person sichergestellt sein kann. Somit wird die vertrauenswürdige Disposition der anderen Person vor dem Hintergrund des verfolgten Interesses von instrumenteller Bedeutung. Die Vertrauenswürdigkeit wird unter diesem Aspekt jedoch zu einem verlässlichen Mittel zum Zweck nivelliert. Sie ist hier als kommensurabel mit einer bloß kausalen Disposition, etwa einer technischen Apparatur dazu, bei bestimmten Inputs bestimmte Outputs zu erzeugen, anzusehen. Man kann in dieser Perspektive deshalb bestenfalls verständlich machen, was es heißt, sich auf die Handlungsdispositionen einer anderen Person zu verlassen, nicht jedoch, was es heißt, *der anderen Person zu vertrauen*. Zitieren wir an dieser Stelle erneut Lars Hertzberg: „When I trust someone, it is *him* I trust; I do *not trust certain things about him*.“[260]

258Scheffler 2011, 28.

259Der Punkt ist hier aber nicht einmal, dass es sich dabei um ein *eigennütziges* Interesse handeln muss. Vielmehr ist entscheidend, dass es mit Vertrauen in eine Person nicht vereinbar ist, diese als bloßes Mittel für einen Zweck zu betrachten – unabhängig davon, ob es sich um einen persönlichen oder kollektiv geteilten Zweck handelt.

260Hertzberg 1988, 315; Hervorhebung von mir.

In dieser Art des sozialkalkulatorischen Zugriffs kommt die andere Person also lediglich als ein zu berechnender Umstand in den Blick. Was kategorial nicht erfasst wird, ist der *normative Anspruch* der anderen Person als vertrauenswürdig angesehen zu werden, geschweige denn, was es heißt, diesen einzusehen, sprich, die andere Person anzuerkennen.[261] Für diese ist die sozialkalkulierende Perspektive schlichtweg blind. Diese aber ist, wie wir gezeigt haben, wesentlich der Bezugsrahmen für Vertrauen in die andere Person. Dadurch bedeutet einer Person zu vertrauen am Ende zwar *auch*, davon auszugehen, dass die andere Person vertrauenswürdig ist. Im Gegensatz zum sozialkalkulierenden Fall hat dies jedoch kategorial zweifach einen ganz anderen Hintergrund: Zum einen ist dies wesentlich als Antwort auf einen normativen Anspruch zu verstehen. Deswegen kann mangelndes Vertrauen als mangelnde Anerkennung erfahren werden, und etwa in persönlicher Verletzung resultieren. Wie wir unten noch genauer zeigen werden hat dies auch wesentlich einen Einfluss auf die Beziehung der beiden Personen zueinander. Zweitens kommt die Vertrauenswürdigkeit der anderen Person nicht einfach als eine bloße Disposition in den Blick. Unsere Erörterung der evaluativen Dimension im Vertrauen hat gezeigt, dass die Vertrauenswürdigkeit des Vertrauensnehmers für den Vertrauensgeber vielmehr eine wertzuschätzende bzw. lobenswerte *Charaktereigenschaft* darstellt, die gerade deshalb lobenswert ist, da die andere Person aus *freier* Einsicht in das normativ Gebotene handelt. Vertrauenswürdigkeit wird deshalb nicht einfach als *kausale* Disposition verstanden, sondern gerade als ein Ausdruck der *Selbstbestimmung* der anderen Person.

6.3.4 Diskussion zweier Einwände

Insgesamt scheint also vieles für die Anerkennungsthese des Vertrauens zu sprechen. Sie vermag es insgesamt ein adäquates und kohärentes Bild unseres gewöhnlichen Verständnisses interpersonellen Vertrauens, was es heißt *jemandem* zu vertrauen, zu zeichnen. Bevor wir dieses Bild in den kommenden Kapiteln erweitern und dessen explanatorisches Potenzial weiter entfalten werden, sollen an dieser Stelle noch auf zwei mögliche Einwände gegen die Anerkennungstheorie diskutiert werden.

So mag man erstens einwenden, dass sich unsere bisherige Diskussion paradigmatisch am Beispiel der Zeugenschaft, genauer am Sprechakt der Mitteilung orientiert hat. In diesem Fall

261 Auch wenn sich der Sozialkalkulierende irrezubel dazu verhält, wenn auch negativ. Siehe diesen Gedanken bei Lagerspetz: „there is no such thingas taking the situation in a neutral way“ (1998, 152).

scheint besonders eindrücklich zu sein, dass Vertrauen eine Form von Anerkennung darstellt, da, der Sprechakt, wie gesagt, selber üblicherweise mit einem impliziten Anerkennungsgesuch einhergeht: Wer jemanden etwas mitteilt („Ich gebe dir mein Wort: p ist der Fall“) möchte (abgesehen vielleicht besonderer Umstände) *ernst genommen* werden, und dass einem aufgrund dessen geglaubt wird. Die Frage bleibt damit jedoch offen, wie Fälle von Vertrauen einzuordnen sind, denen kein solch expliziter Sprechakt vorangeht. Denken wir etwa an den obigen Fall, in dem ich meinem Nachbarn vertraue, dass er nicht das Paket öffnet, das er für mich angenommen hat. Sollte man auch in diesem Fall davon sprechen, dass ich meinen Mitbewohner anerkenne?

Prinzipiell gilt, dass mit der Anerkennungstheorie überhaupt nicht ausgeschlossen ist, dass auch solche Fälle als Fall genuinen interpersonellen Vertrauens aufzufassen sind. Aus der Anerkennungstheorie folgt immerhin nicht, dass die normative Erwartung des Vertrauenspartner *kommuniziert* werden muss. Wir können laut Anerkennungstheorie bereits dann von interpersonellem Vertrauen sprechen, sofern der Vertrauende nur *annimmt*, dass eine (auch unausgesprochene) normative Erwartung ihm gegenüber besteht, den Vertrauenspartner als vertrauenswürdig anzuerkennen. Dies ist zudem in einem Fall der vorliegenden Art auch ganz naheliegend: Es scheint zumindest nichts dagegen zu sprechen, anzunehmen, dass mein Nachbar in dem gegebenen Beispiel einen impliziten normativen Anspruch mir gegenüber haben kann, dass ich ihn als vertrauenswürdig ansehe.

Der Zeugenschaftsfall diente so gesehen für unsere Diskussion als orientierendes Leitbild zum Verständnis von Vertrauen. Auf dem Gebiet der Zeugenschaft hat man einen Fall vorliegen, in dem es besonders eindrücklich um personale Anerkennung geht, da dies als inhärentes Ziel des Sprechaktes der Mitteilung erkennbar wird. Aus der Betrachtung dieses Falles lässt sich also besonders klar die anerkennungstheoretische Konzeption interpersonellen Vertrauens gewinnen, die verspricht, unser gewöhnliches Verständnis interpersonellen Vertrauens angemessen und kohärent wiederzugeben. In diesem Sinne hat die paradigmatische Orientierung am Zeugenschaftsfall einen bloß methodischen Zweck, mit dem die tatsächlich Vielfalt an Vertrauensgestalten nicht negiert werden soll.

Davon abgesehen mag es zwar auch Grenzfälle geben, in denen nicht ohne Weiteres klar ist, ob wir es hier mit interpersonellen Vertrauen im anerkennungstheoretischen Sinne oder einem bloßen sozialkalkulierenden Sich-Verlassen zu tun haben. Wie sind etwa Fälle einzuordnen, in denen man einer Regierung oder einem kollektiven Akteur vertraut?[262] Handelt es sich hier um eine

262Zur Frage nach dem Vertrauen in kollektive Akteure siehe: Hawley 2017.

Anerkennungsbeziehung, oder sollte man schlichtweg von einer bestimmten Form des Sich-Verlassen sprechen? Diese Fragen kann ich an dieser Stelle nicht adressieren. In jedem Fall sprechen diese aber als solche nicht gegen die Anerkennungstheorie. Sie erfragen vielmehr bloß deren *Reichweite*, also in welchen konkreten Fällen es noch sinnvoll ist, diese als Manifestationen interpersonellen Vertrauens im Sinne der Anerkennungstheorie zu charakterisieren. Damit stellen sie keinen Einwand gegen die Theorie als solche dar.

Diese bisherigen Erläuterungen zum Vertrauen als Anerkennung provozieren zudem einen zweiten kritischen Einwand: Wenn die Basis für Vertrauen hier von rein evidenziellen Überlegungen abzugrenzen ist, stellt sich die Frage, ob wir vollkommen jenseits epistemischer Überlegungen vertrauen. Ist es dann also letztlich entweder nur „up to us" oder eine rein praktische Überlegung, und keine epistemische, ob wir einer Person vertrauen oder nicht? Wenn dies so wäre, drohte Vertrauen jedoch eine epistemisch vollkommen *blinde* Haltung darzustellen, die sich in keinster Form etwa an Überlegungen orientiert, die für oder gegen die Vertrauenswürdigkeit einer Person sprechen.[263] Tatsächlich aber scheinen Gründe, die für oder gegen die Vertrauenswürdigkeit einer Person sprechen, eine entscheidende Rolle zu spielen bei der Frage, wem wir letztendlich vertrauen und wem nicht. Wir vertrauen zumindest nicht *vollkommen* unter Absehung solcher epistemischen Erwägungen.

In der Tat stellt sich damit ein tieferliegendes Problem, das wir im vierten Teil dieser Arbeit ausführlicher angehen und diskutieren werden. An dieser Stelle können wir aber bereits festhalten, dass vieles dafür spricht Anerkennung als die angemessene *Kategorie* zur Beschreibung interpersonellen Vertrauens anzusehen. Beim Vertrauen in eine Person handelt es sich *kategorial* gesehen um eine anerkennende Haltung, und kann insofern nicht auf die *evidential view*, wie Moran es nannte, reduziert werden. Dies bedeutet aber gerade nicht, dass Vertrauen vollkommen *jenseits* von Evidenzen, die für oder gegen die Vertrauenswürdigkeit sprechen zu verorten ist. Ganz im Gegenteil werden wir im vierten Teil noch genauer zeigen, dass personale Anerkennung gerade notwendigerweise mit solchen Gründen zusammenhängt. Die theoretische Herausforderung besteht aber darin, beides, Anerkennung und evidenzielle Gründe, auf die *richtige Weise* zusammen zu denken.

263 Im Bereich der Zeugenschaftsdiskussion stellt sich dieses Problem in noch einmal zugespitzter Form, da es hier um die Frage geht, wie man auf der Basis von Mitteilungen anderer zu Wissen gelangen kann. Wenn Vertrauen wesentlich die Antwort auf eine normative Erwartung darstellt, scheint es sich dabei jedoch um ein allein praktisches Phänomen zu handeln. Wie kann ein solches *praktisches* Phänomen aber eine *theoretische* Überzeugung fundieren?

6.4 Vertrauen und die Beziehung der beiden Vertrauenspartner

Die Anerkennungstheorie verspricht insgesamt einen theoretischen Zugriff auf das Vertrauensphänomen zu eröffnen, das weitergehendes explikatorisches Potential in sich trägt. Dies werden wir in diesem Unterkapitel anhand des Zusammenhanges zwischen Vertrauen und der Beziehung zwischen den beiden Vertrauenspartnern herausstellen. Mithilfe der Anerkennungstheorie kann ein plausibles Bild des genauen Stellenwert von Vertrauen im Kontext der Relation zwischen Vertrauensgeber und Vertrauensnehmer gezeichnet werden.

Dieser Stellenwert lässt sich in Hinblick auf zwei Gesichtspunkte beleuchten: Zum einen bildet nämlich die spezifische Art der Beziehung zwischen den beiden Personen einen entscheidenden, auch normativ ausschlaggebenden, Hintergrund für Vertrauen und Misstrauen. Die Signifikanz von Vertrauen und Misstrauen in eine Person ist auf bestimmte Weise von der Beschaffenheit der Beziehung zu dieser abhängig. Damit hängt zum anderen zusammen, dass Vertrauen und Misstrauen einen beziehungs*gestaltenden* Zug haben. Ob man sich gegenüber einer Person vertrauend oder misstrauend verhält, hat einen Einfluss auf die Entwicklung des Verhältnisses zu dieser Person. Wie genau Vertrauen und die Beziehungsart der Vertrauenspartner zusammen stehen, soll im Folgenden beleuchtet werden.

Dabei ist zunächst darauf aufmerksam zu machen, dass unser oben entwickeltes Verständnis von Vertrauen als Anerkennung die konkrete Natur der Beziehung zwischen den Vertrauenspartnern offen lässt. In der aktuellen Vertrauensliteratur gibt es demgegenüber durchaus die Tendenz, sich nur auf bestimmte Arten von Beziehungsformen als Inbegriffe von Vertrauensbeziehungen zu fokussieren. Insbesondere sehen einige Autor:innen Vertrauen als ein Phänomen an, das primär in vergleichsweise affektiven, persönlichen Beziehungen zu verorten ist.[264] Dieses Vorgehen mag insofern seine methodischen Vorzüge haben, als sich bestimmte Aspekte von Vertrauen im Kontext solcher personaler Verhältnisse besonders klar ablesen lassen. Problematisch ist es aber dann, wenn Vertrauensrelationen mit solchen intimen Beziehungen *gleichgesetzt* werden. Dies stellt eine irritierende Einengung des Vertrauensbegriffes dar, die zumindest gerechtfertigt werden müsste. Immerhin vertrauen wir anderen Personen auch im Rahmen förmlicher, anonymer oder flüchtiger Begegnungen, etwa wenn wir Vertrauen in eine Bürokollegin, die Zahnärztin oder andere Verkehrsteilnehmer haben. Die vorgeschlagene Anerkennungstheorie des Vertrauens verspricht jedoch weit genug gefasst zu sein, um dem Spektrum an möglichen Beziehungskontexten von Vertrauen, von den intimsten bis zu den eher

264Vgl. Hartmann 2011, 27.

formalen Beziehungen, gerecht zu werden. Auch dem Fremden auf der Straße nach dem Weg zu fragen und ihm zu glauben ist eine Form ihn ernst zu nehmen und anzuerkennen.

Den hier relevanten Unterschied zwischen intimen und unpersönlichen Beziehungsformen, der auch den je unterschiedlichen Stellenwert von Vertrauen begründet, werden wir im Folgenden anhand der normativen Struktur derselben festmachen. Es gäbe natürlich noch andere Kriterien, um zwischen persönlichen und nicht-persönlichen Beziehungsformen zu unterscheiden.[265] Für unser Anliegen, den spezifischen Stellenwert von Vertrauen und Misstrauen herauszustellen, wird es sich aber als fruchtbar erweisen, auf die für persönliche Beziehungen charakteristische Verbindlichkeitsbeschaffenheit abzuheben. Um dies genauer zu erläutern gehen wir zunächst auf das normative Gefüge einer prototypischen persönlichen Beziehungsform, nämlich der Freundschaft, etwas genauer ein. Anschließend erläutern wir welche Auswirkung dieses für die Relevanz von Vertrauen und Misstrauen in diesem Kontext hat.

Freundschaften werden üblicherweise in einem ersten Schritt als Beziehungen wechselseitigen Wohlwollens charakterisiert.[266] Dies allein wäre aber eine zu kurz greifende Erläuterung, unter anderem deshalb, da somit Freundschaften nicht von einer Beziehung wechselseitiger Gönnerschaft unterschieden wären. Freundschaft ist aber allein insofern mehr als bloße Gönnerschaft, als sie mit einer gewissen *normativen Verbindlichkeit* einhergeht: Als Freund ist man zu bestimmten Handlungen und Haltungen verpflichtet.

Mit einem Ausdruck von R. Jay Wallace sind Freundschaften also durch das strukturiert, was er *freundschaftskonstitutiven Normen* nennt.[267] Darunter sind diejenigen normativen Anforderungen zu verstehen, die man als Freund dem anderen gegenüber hat, und die Freunde auch reziprok voneinander erwarten. Sie spezifizieren, welche Haltungen man dem Freund gegenüber haben und zeigen, und welche Handlungen man ihm gegenüber ausführen sollte. Als Freund gibt es bestimmte Dinge, die man zu tun oder zu lassen hat: So sollten Freunde in Notsituationen füreinander einstehen und dem anderen Mut zusprechen, sie sollten den Kontakt zueinander pflegen, sie sollten Interesse am Leben des Anderen haben, einander im Krankenhaus besuchen etc. Echte Freunde orientieren sich dabei nicht nur in ihrem je eigenem Handeln an diesen Normen, sie erwarten auch vom Anderen, dass er diese Normen berücksichtigt. Wenn diese Normen gebrochen werden, kann dies wiederum als Anlass genommen werden, dem Freund einen

265Axel Honneth und Beate Rössler machen den Unterschied beispielsweise daran fest, dass es in persönlichen Beziehungsformen primär um die Beziehung zu einer *bestimmten* anderen Person geht, während man es in formalen Beziehungen mit prinzipiell ersetzbaren Personen überhaupt zu tun hat (Vgl. Honneth/Rössler 2008, 10f.).

266So etwa in: Helm 2017.

267Wallace 2011, 358.

Vorwurf zu machen. Dabei ist zu beachten, dass es sich auch bei solchen Normen nicht um deontische Normen handelt: Man hat kein *Anspruchsrecht* gegenüber dem Freund, dass dieser sich wie ein Freund verhält. Man fordert es vielmehr von der anderen Person *als* Freund ein. Diese normative Anforderung erhält ihre Geltung also im Lichte der besonderen Art der Beziehung zur anderen Person. Wenn man einem Freund nicht-freundschaftliches Verhalten vorhält, appelliert man deswegen auch typischerweise an das besondere Verhältnis, das man zueinander einnimmt: „Ich dachte wir sind Freunde!".

Was genau Freunde wechselseitig voneinander erwarten orientiert sich natürlich grundlegend an kulturell grob vorgegebenen Vorstellungen darüber, was eine Freundschaft idealtypischerweise ausmacht. Darüber hinaus unterscheiden sich Freundschaften in den genaueren Details ihrer Verbindlichkeitsreichweite. Manchen Freunden mag es beispielsweise ganz recht sein, wenn man voneinander nicht all zu viel erwarten muss, innerhalb anderer Freundschaften haben die Beteiligten relativ weitreichende normative Erwartungen aneinander.[268] Diese spezifischen normativen Anforderungen im Rahmen einer konkreten Freundschaft können wiederum zum Teil das Ergebnis expliziter Abmachungen sein, zum anderen Teil mögen sie aber auch auf sich stillschweigend eingespielten Routinen zwischen den Freunden beruhen.[269]

Ein interessanter Punkt, auf den Wallace hinweist, ist, dass die Frage ob man einer solchen freundschaftskonstitutiven Norm nachkommt oder nicht immer auch eine *expressive Signifikanz* hat. Einer solchen Norm zu folgen, kommt demnach nicht einfach dem bloßen Einsehen in die Notwendigkeit von *Regeln* gleich, etwa so wie man beim Schachspiel die Spielregeln beachtet. Es stimmt zwar, dass Freunde nicht einfach „*irgendwie*" wohlwollend aufeinander eingestellt sind, insofern freundschaftskonstitutive Normen *spezifizieren*, was genau im Rahmen der Freundschaft verlangt ist. Der *Grund weshalb* man sich an diesen Normen orientiert, liegt jedoch nicht darin, dass diese als notwendige Spielregeln anerkannt werden, sondern darin, dass einem persönlich das Wohlergehen des Anderen wichtig ist: „A genuine friend will visit you in the hospital when you are convalescing not merely out of a cold and calculating recognition that friendship requires this of them, but *out of a concern* for your well-being".[270] Vor diesem Hintergrund erhält das Befolgen oder Nicht-Befolgen freundschaftskonstituiver Normen eine besondere *expressive Signifikanz*. Das heißt, dass das Nichtnachkommen der Verbindlichkeiten gegenüber einem Freund *prima facie* immer auch auch *Ausdruck* einer mangelnden Sorge um die andere Person ist. Es drückt damit

268Siehe hierzu auch: Löschke 2017.
269Diesen Punkt haben wir oben (Kap. 5.3) in Bezug auf implizite, praxisgebundene *commitments* erwähnt.
270Wallace 2011, 357.

auch mangelnde Wertschätzung der Beziehung zu dieser Person aus.

Was nun insbesondere zur Freundschaft dazugehört und auch reziprok voneinander erwartet wird, ist, dass man sich gegenseitig ernst nimmt, oder zumindest eine Bereitschaft dazu an den Tag legt, die weiter geht, als das gegenüber Fremden der Fall wäre. Entsprechend gehört insofern zur Freundschaft auch, dass man sich stärker gegenseitig vertraut, als man dies gegenüber Fremden tun würde.[271] Als guter Freund zeigt man deshalb eine höhere *Bereitschaft* dem anderen zu vertrauen. Für enge Beziehungen wie Freundschaftsbeziehungen ist Vertrauen, sowie übrigens auch Vertrauenswürdigkeit, eine geradezu konstitutive normative Anforderung. Judith Baker schildert einen im Vertrauensdiskurs prominent diskutierten Fall, der diese freundschaftsgebundene Anforderung zu vertrauen, illustriert. Angenommen eine Freundin von mir sei wegen Geheimnisverrates von der Regierung angeklagt, versicherte mir jedoch, entgegen vieler Beweise, die für ihre Schuld sprechen, dass sie unschuldig sei. In diesem Fall, so legt Baker hier nahe, komme ich als guter Freund der normativen Anforderung nach, eine Bereitschaft dazu zu haben, ihr zu vertrauen: „I believe [my friend] is innocent. [...] I believe there is an explanation for the alleged facts, for the accusation, which will clear it all up. In advance of hearing the case, I am prepared to believe that there is such an explanation. I am biased in favour of my friend, in favour of her innocence. To put it another way, *I am committed to her being innocent*.“[272]

Wie eben dargestellt, hat die Frage, ob man freundschaftskonstitutive Normen folgt oder nicht immer auch eine expressive Signifikanz. Dies gilt auch auf dem Feld von Vertrauen und Misstrauen. Eine hohe Bereitschaft dazu einem Freund zu vertrauen, drückt immer *auch* aus, dass man sich an den freundschaftskonstitutiven Normen orientiert, und hat damit im Wallace‘schen Sinne auch eine besondere expressive Bedeutung. Eine hohe Bereitschaft zu Vertrauen *bedeutet*, dass man die Freundschaft zur anderen Person hochhält. Vertrauen drückt in diesem Zusammenhang also immer auch eine Bestätigung der freundschaftlichen Beziehung aus. Umgekehrt kann eine Neigung zu Misstrauen gegenüber einem Freund eine mangelnde Wertschätzung der Beziehung zur anderen Person ausdrücken.

Eine Bereitschaft zu Vertrauen oder Misstrauen drückt also immer auch etwas über den Stellenwert, den das Verhältnis zur anderen Person hat aus. Damit zusammen hängt, dass Vertrauen und Misstrauen zugleich auch einen wesentlich *beziehungsgestaltenden* Zug haben.[273] Vertrauen bestärkt eine intime Beziehung, Misstrauen kann eine intime Beziehung irritieren oder gar

271 Auch Cogley (2012) geht von einer solchen Norm in persönlichen Beziehungen aus, allerdings ohne sie mit Hilfe des Konzeptes der Anerkennung zu erläutern.

272 Baker 1987, 3.

273 Vgl. Harding 2011, der diesen Gedanken aber auch nicht mit Bezug auf eine Anerkennungstheorie erläutert.

zerstören. Vertrauen in eine noch fremde Person kann zudem eine prinzipielle Bereitschaft ausdrücken, die Beziehung zu dieser intimer und verbindlicher zu gestalten. Oben hatten wir diesen Gedanken mit einer Bemerkung von Holton bereits einmal angedeutet: „I trust you; in so doing our relationship moves a little further forward. This can itself be something I value.“[274] Vertrauen und Misstrauen stellen also auch immer eine Form von beziehungsgestaltendem Verhalten dar. In diesem Sinne bekommt Vertrauen auch eine andere *Relevanz* je nach Kontext der Beziehung. Ob man vertraut oder nicht vertraut, hat nämlich Auswirkung auf die Beziehung zur anderen Person. Offenes Misstrauen etwa einem Freund gegenüber bedeutet, den freundschaftskonstitutiven Normen nicht zu genügen, und kann insofern die freundschaftliche Beziehung infrage stellen. In diesem Sinne[275] ist Vertrauen in diesem Kontext besonders relevant.

6.5 Zusätzliche Bemerkungen

Ausgehend von der Anerkennungstheorie des Vertrauens und den Ausführungen zum Zusammenhang zur Beziehung der beiden Vertrauenspartner können noch weitere Fragen, die sich in Bezug auf Vertrauen stellen beleuchtet werden. An dieser Stelle möchte ich nur zwei Fragen anreißen, die wiederholt in der Vertrauensliteratur aufgeworfen wurden. Die erste betrifft das normative Eigengewicht von Vertrauen, die zweite noch einmal die Phänomenologie der reaktiven Einstellungen im Falle enttäuschten Vertrauens. Die folgenden Überlegungen sind jedoch recht tentativ und sollen nur einen Ausblick zur Behandlung von in der Literatur wiederholt aufgeworfenen Fragen ausweisen.

1) In der Literatur wird vereinzelt die Frage aufgeworfen, ob Vertrauen als solches eine gewisses normatives Gewicht mitsichbringt.[276] Zwar hatten wir bereits festgestellt, dass Vertrauen wesentlich im Kontext normativer Anforderungen situiert ist. Damit blieb aber die Frage offen, ob Vertrauen als solches eine *zusätzliche* normative Anforderung an den Vertrauensnehmer mitsichbringt, oder ob es bloß auf ohnehin vertrauensunabhängig bestehenden normativen Erwartungen aufsitzt, selber

274Holton 1994, 68.

275Ob man vertraut oder misstraut kann natürlich auch in vielen anderen, etwa in instrumenteller Hinsicht, extrem relevant sein. Gemeint ist hier allein die Relevanz von Vertrauen für die Beziehung zur einer konkreten anderen Person. Der Punkt ist, dass man diese Art der Relevanz von Vertrauen und Misstrauen nur erläutern kann, wenn man Vertrauen als Anerkennung versteht. Erst vor diesem Hintergrund erklärt sich der beziehungsinterne Stellenwert von Vertrauen.

276Etwa von Collin O‘Neil (2017).

aber normativ leer ist.

Ein Ausgangspunkt, um über diese Frage nachzudenken, kann die Feststellung sein, dass es sich bei Vertrauen um eine Form von Anerkennung der Vertrauenswürdigkeit der Vertrauensperson handelt. Die Frage, die sich dann stellt ist, ob die Tatsache, dass der Vertrauensnehmer anerkannt wird, für jenen zusätzlich normativ ausschlaggebend ist. Ausgehend von der oben herausgestellten Bedeutung, die Vertrauen für die Gestaltung einer Beziehung hat, können wir einen eigenen Ausblick auf diese Frage geben. Vertrauen ist demnach insofern in sich normativ relevant als es (zumindest unter bestimmten Umständen) den Versuch zur Vertiefung oder die Bestätigung der Beziehung zum Vertrauensnehmer bedeuten kann. Vertrauen kann in diesem Sinne auch etwas Appellierendes an den Vertrauenspartner haben. Sich dem gegenüber als nicht vertrauenswürdig zu erweisen stellt insofern immer *auch* einen negativen Respons auf einen solchen Versuch zur Gestaltung der Beziehung zur anderen Person dar. Dies mag in sich selbst ein gewisses normatives Gewicht in den praktischen Überlegungen des Vertrauensnehmers darstellen.

2) Vertrauen zu brechen, stellt einen typischen Anlass für reaktive Einstellungen oder Emotionen dar. Wie oben dargestellt, werden diese üblicherweise als Ausdruck der Verletzung einer normativen Erwartung an den Vertrauenspartner gesehen. Einige Autor:innen stellen aber heraus, dass die Verletzung von Vertrauen zu ganz *spezifischen* reaktiven Emotionen Anlass gebe.[277] Stephen Darwall etwa argumentiert, dass reaktive Einstellungen wie Vorwürfe und auch typischerweise damit einhergehende Emotionen wie Wut, nicht der eigentliche Ausdruck enttäuschten Vertrauens seien. Sie sind womöglich Ausdruck einer berechtigten normativen Erwartung an den Vertrauenspartner, die den Kontext bilden, vor dem man einer Person vertrauen kann. Sie seien jedoch nicht Ausdruck der Enttäuschung des Vertrauens *selber*. Demgegenüber schlägt Darwall vor, dass sich enttäuschtes Vertrauen in, etwas tentativ formuliert, persönlicheren emotionalen Reaktionen ausdrücke. Diese lässt aber aber weitestgehend unanalysiert und zählt lediglich einige unsortierte Umschreibungen auf: Eine Verletzung von Vertrauen „appropriately give[s] rise to distinctive attitudes or feelings, such as feeling letting down, personal disappointment, or personal hurt.“[278]

Es scheint, dass sich das Charakteristische von Vertrauen (im Gegensatz zu einer bloß sozial kalkulierenden Einstellung) *auch* symptomatisch in seiner spezifischen Art der Enttäuschung zeigt. Enttäuschtes Vertrauen besteht demnach weder in der Enttäuschung einer epistemischen

277 O‘neil 2016, 78. Siehe auch: Cogley 2012.
278 Darwall 2017, 35.

Erwartung, noch in der Enttäuschung einer normativen Anforderung, die die Vertrauensperson gegenüber dem Vertrauenden hatte. Sie besteht vielmehr in der *Enttäuschung des Vertrauenden*. Um diese spezifische Form der Enttäuschung aber klar von den beiden anderen abzugrenzen, scheint es mir hilfreich, auf die Anerkennungstheorie zurückzugreifen. Die Enttäuschung des Vertrauenden mag demnach als *Enttäuschung der Anerkennung* des Vertrauenden in den Vertrauenspartner verstanden werden.

6.6 Zwischenfazit: Das explikatorische Potenzial der Anerkennungstheorie

Die in diesem Kapitel entwickelte Anerkennungstheorie des Vertrauens vermag es somit insgesamt auf den Begriff zu bringen, was es im Kern heißt einer Person zu vertrauen. Im Einklang damit vermag sie es auch bestimmte Phänomene unserer Vertrauenspraxis erklären zu können. Zusammenfassend sollen die folgende Vorzüge der Anerkennungstheorie noch einmal zusammengestellt werden:

1) Die Anerkennungstheorie des Vertrauens umgeht das Problem der Sozialkalkulation, das sich für die Standardauffassung des Vertrauens stellte. Sie zielt damit auf den eigentlichen Kern dessen, was es heißt jemandem zu vertrauen und schließt damit aus, dass Vertrauen als bloß sozialkalkulatorische Haltung verstanden werden könnte. Anders ausgedrückt scheint mit der Anerkennungstheorie, im Gegensatz zur Standardauffassung, ein klarer Kontrast zwischen Vertrauen und einem bloßen Sich-Verlassen gezogen werden zu können.

2) Die Anerkennungstheorie kann zusätzliche phänomenologische Besonderheiten von Vertrauen erklären. Dazu gehört zum einen der Umstand, dass aus der Sicht des Vertrauenspartners Vertrauen als Erfahrung der Anerkennung, bzw. mangelndes Vertrauen unter Umständen als verletzend empfunden werden kann.

3) Zudem stellt die Anerkennungstheorie einen adäquaten Ausgangspunkt dar, um den Stellenwert von Vertrauen und Misstrauen zur Gestaltung von personellen Beziehungen zu beleuchten. Dabei ist sie weit genug formuliert, sodass sie Vertrauen nicht zu einem Phänomen einengt, dass es nur in besonders intimen Beziehungen geben könne. Vielmehr umspannt sie die breite Vielfalt an Vertrauensbeziehungen, die wir vorfinden, die etwa auch anonymer, flüchtiger oder formeller Natur sein können.

Damit werden wir vorerst die Betrachtung der interpersonellen Seite von Vertrauen abschließen. Wir werden auf diese im vierten Teil dieser Arbeit noch einmal zurückkommen, wenn es detaillierter darum gehen wird zu zeigen, wie das hier gewonnene anerkennungstheoretische Verständnis von Vertrauen mit der epistemischen Struktur adäquat zusammen gedacht werden kann. Dazu jedoch müssen wir zunächst klären, wie die epistemische Struktur von Vertrauen überhaupt beschaffen ist. Um was für eine Art von Haltung handelt es sich bei Vertrauen aus epistemischer Sicht? Und in welchem begrifflichen Zusammenhang stehen Vertrauen und epistemische Gründe? Diese Fragen werden wir im folgenden dritten Teil dieser Arbeit angehen.

III. Teil: Zur epistemischen Grammatik von Vertrauen

7. Die epistemische Struktur von Vertrauen

7.1 Einleitung: Die Relation zwischen Vertrauen und epistemischen Gründen

Die epistemische Seite von Vertrauen stellt sich als ein mehrfach theoretisch herausforderndes Themengebiet dar. Naheliegend scheint zunächst, dass Vertrauen, sofern es nicht einfach als willkürlich oder a-rational verstanden werden soll, in einem Zusammenhang mit der epistemischen Begründung seiner Angemessenheit stehen muss. Immerhin scheint eine Vertrauende ihr Vertrauen in eine Person P nicht einfach als eine vollkommen arbiträr zustande gekommene Haltung zu verstehen, so als ob sie P ebenso gut auch misstrauen könnte. Insofern sie P vertraut, geht sie vielmehr davon aus, dass P auch vertrauenswürdig *ist*. Wenn dies jedoch stimmt, scheint sie Gründe dafür annehmen zu müssen, weshalb sie den anderen als vertrauenswürdig einschätzt. Zumindest ihrem Anspruch nach gibt es aus der Sicht der Vertrauenden also etwas, was *dafür spricht*, dass die andere Person tatsächlich vertrauenswürdig *ist*.

Diese kurze Argumentationsskizze führt schnell vor Augen, dass Vertrauen in einem noch unklaren begrifflichen Zusammenhang zu epistemischen Gründen stehen muss. Etwas spricht für die Vertrauenswürdigkeit der Person, der ich vertraue, das heißt, es gibt *Gründe* dafür. Bei genauerem Hinsehen stellt sich dieser Zusammenhang zwischen Vertrauen und Gründen aber als äußert problembehaftet und erläuterungsbedürftig dar. Ein großer Teil der philosophischen Diskussion über Vertrauen bezieht sich auf ebendiesen Zusammenhang. Vertrauen und seine epistemischen Gründe hängen nämlich nicht einfach *irgendwie* zusammen, sondern auf eine bestimmte Art und Weise, die für Vertrauen gerade charakteristisch ist. In diesem dritten Teil der Arbeit wird es um diesen Zusammenhang zwischen Vertrauen und seinen epistemischen Gründen gehen. Er soll in verschiedener Hinsicht erläutert werden.

Für die folgende Diskussion in diesem und im nächsten Kapitel der Arbeit stehen vor allen Dingen zwei Hinsichten im Fokus, die einer Erläuterung bedürfen, wobei diese beiden Aspekte m.E. in der gegenwärtigen Literatur nicht immer sauber voneinander getrennt werden.

1) Zunächst einmal ist zu zeigen, dass sich Vertrauen aus trivialen begrifflichen Gründen mit dem aktualen und ernsthaften Erwägen von epistemischen Gründen ausschließt. Man kann schließlich nicht vertrauen und *gleichzeitig* das Vertrauen *ernsthaft* nach seinen Gründen befragen. In diesem Sinne ist Vertrauen qua seiner grundlegenden Grammatik mit Zweifeln und Deliberation nicht vereinbar. Diesen Punkt werden wir am Ende dieses 7. Kapitels adressieren.

2) Der erste Punkt widerspricht nicht der Feststellung, dass wir zuweilen erst *überlegen*, bevor wir uns dazu *entscheiden* zu vertrauen. In der Literatur wurde jedoch zuweilen festgehalten, dass dieser Modus expliziten Entscheidens, zumindest in bestimmten Kontexten, einer vollwertigen vertrauenden Haltung widerspricht oder gar Vertrauen untergraben kann. Dieser Aspekt spricht gewissermaßen den bewusstseinstheoretischen Modus von Vertrauen und seinen Gründen an und soll im 8. Kapitel näher beleuchtet werden.

Diese beiden Punkte gilt es zudem von einem dritten Feld abzugrenzen, auf dem das Verhältnis zwischen Vertrauen und epistemischen Gründen erläuterungsbedürftig ist. Dieser bezieht sich speziell auf interpersonelles Vertrauen und den Zusammenhang zwischen Anerkennung und epistemischen Gründen. Dieses Thema an der Schnittstelle der interpersonellen und epistemischen Seite von Vertrauen werden wir im 9. Kapitel behandeln.

In diesem Kapitel werden wir im Detail zunächst folgendermaßen vorgehen. Zunächst werden wir noch etwas grundlegender auf die epistemische Struktur von Vertrauen eingehen. Wir werden zunächst dafür argumentieren, dass Vertrauen überhaupt einen kognitiven Anteil hat, sprich, dass Vertrauen, in einer bestimmten Definition des Begriffes, eine Überzeugung beinhaltet (7.2). Daran anknüpfend werden wir detaillierter aufzeigen, dass Vertrauen begrifflich mit einem Anspruch auf epistemische Angemessenheit einhergeht, und folglich in einer Verbindung zu epistemischen Gründen steht (7.3). Zuletzt werden wir nachweisen, dass Vertrauen als festgelegte Haltung in einem bestimmten trivialen Sinne mit der bewussten Abwägung von Gründen unvereinbar ist (7.4).

Es sei noch betont, dass die folgende Diskussion nicht auf interpersonelles Vertrauen im Speziellen beschränkt ist, sondern sich auf Vertrauen im weiteren generischen Sinne bezieht. Erst im anschließenden vierten Teil dieser Arbeit legen wir uns die Frage vor, inwiefern die epistemische Struktur von Vertrauen im Allgemeinen mit den Besonderheiten interpersonellen Vertrauens in Einklang gebracht werden kann.

7.2 Der Kognitivismus des Vertrauens

Bereits im einleitenden Teil dieser Arbeit haben wir darauf hingewiesen, dass es sich bei einer

vertrauenden Haltung grundlegend um eine zuversichtliche Haltung handelt. Insofern man vertraut, geht man nicht davon aus, dass der mögliche Risikofall eintreten wird, sondern dass die Sache gut ausgehen wird. Im Bereich interpersonellen Vertrauens haben wir zudem gezeigt, dass sich dieser Optimismus auch auf die Vertrauenswürdigkeit der anderen Person erstreckt. Jemandem zu vertrauen heißt demnach, davon auszugehen, dass der Vertrauensnehmer zumindest bezüglich des gegebenen Vertrauensgegenstandes vertrauenswürdig ist.

Unter Kognitivismus des Vertrauens soll hier die Position verstanden werden, die auf Grundlage dieser Überlegungen den Schluss zieht, dass Vertrauen mit einer epistemischen *Überzeugung* bzw. epistemischen *Erwartung* einhergeht. Insofern man darauf vertraut, dass p passieren wird, ist man der Überzeugung oder erwartet man, dass p passieren wird. Insofern man jemandem vertraut hat man entsprechend die Überzeugung, dass die andere Person vertrauenswürdig ist bzw. man hat die Erwartung, dass sie sich als vertrauenswürdig erweisen wird.[279] Gegen diese These wurden jedoch einige Einwände formuliert. Wir werden uns deshalb detaillierter mit der Diskussion um den kognitiven Status von Vertrauen auseinandersetzen, um am Ende die kognitivistische Lesart gegen übliche Gegenargumente zu behaupten. Wir werden aufzeigen, dass die anti-kognitivistischen Überlegungen vielmehr anzeigen, dass Vertrauen eine *Erwartung spezifischer Art* beinhaltet, die sich von verwandten epistemischen Einstellungen, wie etwa einer festen Gewissheit unterscheidet. Die radikale Schlussfolgerung, dass Vertrauen keine Erwartung beinhalte, ist jedoch ungerechtfertigt.

Eine unmittelbare Anfangsplausibilität gewinnt die kognitivistische Position zunächst durch Betrachtung solcher Aussagen wie „Ich vertraue dir, aber ich glaube nicht, dass du vertrauenswürdig bist“. Diese Äußerung erscheint uns unmittelbar als inkohärent, und geradezu die Qualität eines Mooreschen Satzes[280] zu haben. Würde man etwa sagen, dass man P vertraut ϕ zu tun, aber zugleich die Überzeugung vertreten, dass P in Bezug auf ϕ nicht vertrauenswürdig ist, verhielte man sich schlichtweg widersprüchlich. Dieser Widerspruch scheint demnach zum

279Der kognitivistischen Sichtweise könne etwa folgende Autor:innen zugeordnet werden: Baker (1987); Fricker (2006); Gambetta (2000); Hieronymi (2008); McMyler (2011, 115, 131–41; 2013) und Moran (2005). Ausdrückliche Non-Kognitivisten sind demgegenüber: Faulkner (2007; 2011; 2014); Hawley 2014a; Holton (1994); Jones (1996); Lahno (2001). Siehe auch die Diskussion der Positionen in: Simpson 2018.
Auf beiden Seiten werden die Begriffe Kognitivismus bzw. Non-Kognitivismus nicht immer ganz einheitlich verwendet. Innerhalb des kognitivistischen Lagers gibt es etwa die beiden untergeordneten Positionen, dass Vertrauen entweder durch eine Überzeugung konstituiert ist, oder dass Vertrauen eine Überzeugung zur Konsequenz hat. Wir werden in diesem Kapitel die Position vertreten, dass die Überzeugung konstitutiv zum Vertrauen dazugehört. Im Bereich interpersonellen Vertrauens ist zuweilen auch nicht eindeutig, was die Überzeugung beinhaltet, ob es sich etwa um die Überzeugung handelt, dass die andere Person vertrauenswürdig ist, oder schlichtweg um die Überzeugung, dass sie tun wird, worauf man ihr vertraut. Auf Letzteres hebt etwa die kognitivistische Position von Marusic (2011, 178) oder Hieronymi (2008, 227) ab.

280Siehe Moore 1993, 210.

Ausdruck zu bringen, dass wir Vertrauen als Haltung verstehen, die die Überzeugung in die Vertrauenswürdigkeit der anderen Person inkludiert. Vertrauen ist in diesem Sinne wesentlich auch eine kognitive Haltung.

Vielleicht mag man einwenden, dass dieser Moore'sche Widerspruch doch nur die Unvereinbarkeit von Vertrauen in P und der Überzeugung, dass P *nicht* vertrauenswürdig ist zum Ausdruck bringt.[281] Dies müsse doch nicht zwangsläufig bedeuten, dass Vertrauen die positive Überzeugung in die Vertrauenswürdigkeit des anderen umfasst. Als Replik auf diesen Einwand beachte man aber, dass der folgende abgeschwächte Moore'sche Satz ebenso inkohärent zu sein scheint: „Ich vertraue dir, aber ich weiß nicht, ob du vertrauenswürdig bist". Die Inkohärenz dieser Aussage bringt zum Vorschein, dass mit dem Vertrauen in eine Person Zweifel oder Unklarheit über die Vertrauenswürdigkeit einer Person ausgeschlossen sind. Insofern man vertraut, ist man also überzeugt davon, dass die andere Person vertrauenswürdig ist.[282] Deshalb erscheint die Aussage widersprüchlich.

Ein weiterer Einwand mag an dieser Stelle lauten, dass es auch Fälle zu geben scheint, in denen man sich dafür entscheiden kann zu vertrauen, obwohl man Zweifel hat. Richard Holton etwa diskutiert dies am Beispiel eines „trust circles", einer Übung, die beispielsweise in Theater-Kursen praktiziert wird, und in denen sich eine Person mit verbundenen Augen rücklings fallen und von den Schauspielkolleginnen auffangen lassen muss. In einer solchen Situation, so Holtons Argument, könne man doch durchaus beängstigende Zweifel darüber haben, ob man aufgefangen wird, aber sich dennoch dazu entscheiden zu vertrauen.[283] Kann man hier also seine Zweifel nicht einfach ignorieren oder übergehen, um zu Vertrauen zu gelangen, insbesondere dann, wenn man ein starkes Ziel vor Augen hat, etwa die Stärkung der Gemeinschaft der Gruppe?

Gegen diesen Einwand hat Pamela Hieronymi jedoch argumentiert, dass der von Holton geschilderte Fall kein Beispiel genuinen Vertrauens darstelle, sondern lediglich einen Fall dessen, was sie *entrusting* nennt. Wir kennen diese Kategorie bereits unter anderer Bezeichnung von unserem einleitenden Teil. Dort haben wir zwischen dem bloßen Eingehen eines Risikos bzw. eines *Vertrauensaktes* und einer genuin *vertrauenden Haltung* unterschieden. *Entrusting* nun aber ist als deckungsgleich mit Ersterem anzusehen und bezeichnet also lediglich den Akt, mit dem man sich gegenüber einer anderen Person verletzlich macht, unabhängig davon, ob man dabei optimistisch

281 Siehe dieses Argument in: Holton 1994, 71.

282 In jedem Fall können wir sagen, scheint damit kein vollkommenes Vertrauen zum Ausdruck gebracht worden zu sein. Wir können die kognitivistische These dafür auch einfach graduell formulieren. Demnach lautete sie: Vertrauen kommt in Graden, und je robuster die Überzeugung in die Vertrauenswürdigkeit des anderen, desto stärker ist das Vertrauen.

283 Holton 1994, 63f.

oder pessimistisch eingestellt ist. Es stellt damit eine bloße „decision to rely in the face of doubt“[284] dar. Damit handelt es sich hierbei nicht wirklich um Vertrauen, sondern nur um eine Art Pseudo-Vertrauen, ein Handeln *als ob* man vertraute. Auch wenn wir in solchen Situationen alltagssprachlich die Vokabel „Vertrauen“ verwenden, scheint diese lediglich darauf zu referieren, dass eine Vertrauenshandlung vollzogen werde, nicht aber, dass diese von *tatsächlichem* Vertrauen begleitet war.

Womöglich mag man hier weiter nachhaken wollen und den Einwand einbringen, dass man doch auch in einer solchen Situation wie im „trust-circle“ *sowohl* Zweifel haben *als auch* vertrauen könne, etwa insofern man zwischen beidem hin- und herschwanke. Natürlich aber ist es *psychologisch* möglich zwischen diesen beiden Zuständen zu changieren. Hier geht es jedoch um einen grundlegenderen *begrifflichen* Punkt, der besagt: *insofern* jemand vertraut, zweifelt er nicht an der Vertrauenswürdigkeit der Vertrauensperson. *Insofern* er an der Vertrauenswürdigkeit zweifelt, vertraut er nicht. Im begrifflichen Sinne schließen sich Vertrauen und Zweifel aus. Dies schließt nicht aus, dass man eine Person als zugleich vertrauend und zweifelnd beschreiben mag, etwa um ihre schwankende Haltung oder ihre innere Zerrissenheit zum Ausdruck zu bringen. Ganz im Gegenteil bestätigt dies gerade unseren Punkt: Die Person wäre schließlich gerade insofern *zerrissen*, als Vertrauen und Zweifel begrifflich gesehen im *Widerspruch* zueinander stehen.[285]

Aus dem gleichen Grund lässt sich auch kein non-kognitvistisches Argument dadurch gewinnen, dass man auf vermeintliche Sonderformen von Vertrauen hinweist, in denen man ohne Überzeugung „vertraut“.[286] Zu diesen zählt insbesondere das sogenannte „therapeutische Vertrauen“.[287] Dieses besteht darin, einer anderen Person entgegen der fehlenden Überzeugung in ihre Vertrauenswürdigkeit zu „vertrauen“, mit dem Ziel diese andere Person (kurz oder langfristig) zu einer vertrauenswürdigeren Einstellung zu ermutigen.[288] Tatsächlich jedoch scheint dies kein

284Hieronymi 2008, 9.

285In vergleichbarer Weise unterstreicht Bela Szabados, dass Wittgensteins Bemerkung, dass man seinen eigenen Überzeugungen nicht misstrauen könne als grammatische Aussage zu verstehen sei. Damit ist es weiter zulässig eine Person in einer psychologischen Beschreibung als misstrauisch gegenüber ihren eigenen Überzeugungen auszuweisen (Szabados 1981).

286So etwa Frost-Arnold (2014). Neben dem in diesem Abschnitt erwähnten therapeutischen Vertrauen unterscheidet Frost-Arnold noch zwischen *coping trust* und *corrective trust. Coping trust* bezeichnet dabei die Entscheidung aus pragmatischen Gründen zu „vertrauen“, da dies etwa praktische Planungen einfacher mache. *Corrective trust* ist eine Entscheidung für „Vertrauen“ im Lichte möglicher Vorurteile, die man gegenüber dem Vertrauenspartner hegt und die es auszugleichen gelte. Mit dem gleichen Argument, mit dem ich hier therapeutisches Vertrauen nicht als tatsächliches Vertrauen einordne, gehe ich davon aus, dass auch *coping trust* und *corrective trust* keine Formen genuinen Vertrauens darstellen.

287Siehe die klassische Definition von Howard John Neate Horsborough: „A reliance [of this sort, i.e. one] which aims at increasing the trustworthiness of the person in whom it is placed, may be described as a therapeutic trust“ (1960, 346).

288Siehe auch die Thematisierung in: Faulkner (2007); Holton (1994, 63); Jones (1996, 23); McGeer (2008).

geeignetes Gegenbeispiel darzustellen, um zu beweisen, dass Vertrauen und Überzeugung nicht notwendig zusammenhängen. Denn auch hier haben wir es wiederum allein mit Fällen bloßer Vertrauens*akte* – mit Hieronymi „entrusting" genannt – zu tun. Die Beispiele scheinen demnach lediglich Vertrauensakte zu beschreiben, die gerade keine wirkliche vertrauende Haltung darstellen, sondern deren Witz ja gerade ist, dass die Akteurin nur *so tut als ob*, er vertraute, ohne *wirklich* zu vertrauen. Mehr noch bestätigt der Fall geradezu, dass wir Vertrauen eigentlich kognitiv verstehen. Denn ein ebensolches *begriffliches* Verständnis scheinen die Akteure auch im Falle therapeutischen Vertrauens *vorauszusetzen*: Ebendiese Form von Vertrauen versucht die therapeutisch Vertrauende ja gerade vorzugeben bzw. meint die Vertrauenspartnerin in der therapeutisch Vertrauenden zu sehen. Dies zeigt sich an folgendem Umstand: Würde die therapeutisch „Vertrauende" ihrem „Vertrauenspartner" offen sagen, dass sie nicht an die Vertrauenswürdigkeit der anderen Person glaubt, würde sie auch nicht mehr sinnvoll vorgeben können zu vertrauen, und damit wäre auch Vertrauen im therapeutischen Sinne nicht mehr möglich.

Die bisherigen Arten von Einwänden zielten darauf ab Vertrauen als eine Sache bloßer Entscheidungen auszuweisen und verkennen damit den kognitiven Charakter desselben. Eine andere Gruppe von Einwänden, denen wir uns nun zuwenden werden, erweisen sich demgegenüber eher in der Wortwahl verwirrend. Die Ansätze, um die es geht, argumentieren, dass es sich bei Vertrauen zumindest nicht um eine vollwertige epistemische *Überzeugung* handle. Im Zuge dieser Abgrenzung weisen sie zwar korrekterweise einige Besonderheiten der epistemischen Dimension von Vertrauen nach, etwa inwiefern sich Vertrauen von anderen epistemischen Haltungen, wie Wissen oder Sich-sicher-sein abgrenzt. Daraus aber zu schließen, dass Vertrauen *keinerlei* Überzeugung oder Erwartung beinhalte, ist eine zu starke Schlussfolgerung. Diese beruht zumeist darauf, dass eigene – und teilweise recht idiosynkratische – Begriffe von Überzeugungen oder Erwartungen vorausgesetzt werden, von denen dann eine vertrauende Haltung abgegrenzt werden soll.

Ich werde diesen Punkt am Beispiel von Bernd Lahnos ausführlichem Versuch, Vertrauen von einer Erwartung abzugrenzen, verdeutlichen. Lahnos Ansatz bietet sich hier an, da er erstens in einigen Textstellen recht ausführlich und klar versucht den Unterschied zwischen Vertrauen und einer Erwartung zu ziehen und dabei zweitens meines Erachtens auch einige richtige und in der Literatur geteilte begriffliche Zusammenhänge darstellt. Allerdings scheint es mir ein *non sequitur* auf dieser Grundlage darauf zu schließen, dass der Kognitivismus falsch sein muss.[289]

289Auch Lahno bezieht diese Diskussion nicht auf interpersonelles Vertrauen im Speziellen, sondern auf das, was er Vertrauen-darauf nennt. Dies definiert er als den propositionalen Aspekt von Vertrauen. Um zu erläutern, was er

Kommen wir damit nun zu Lahnos Unterscheidungsversuch. Für Erwartungen und Vertrauen postuliert er zunächst einmal, dass diese einen gemeinsamen Nenner haben: Beide implizieren die *Annahme, dass ein Ereignis p*[290] eintreten wird. Beide bezeichnen einen „geistigen Zustand, in dem man sich auf eine bestimmte Entwicklung einstellt und sein Handeln an dieser Einstellung ausrichtet.“[291] Darüber hinaus ist festzustellen, dass sich Vertrauen durch zwei spezifische Charaktersistika auszeichnet. So bezieht sich Vertrauen erstens stets auf positive zukünftige Ereignisse, die zweitens für die vertrauende Person eine gewisse Relevanz haben.

Diese zwei Unterscheidungsmerkmale zeichnen jedoch noch keinen wesentlichen Unterschied zwischen Vertrauen und Erwartungen. Vielmehr scheinen sie Vertrauen lediglich als eine durch ihren Gegenstandsbereich ausgezeichnete Subkategorie von Erwartung zu charakterisieren: Und zwar eine Erwartung, die sich auf positiv-bedeutsame Ereignisse bezieht. Allein mit dem dritten Unterscheidungskriterium unternimmt Lahno den Versuch Vertrauen von Erwartungen *als solchen* begrifflich abzugrenzen. Vertrauen ist demnach keine besondere Form von Erwartung, sondern eine grundlegend andere Art von Einstellung. Diesbezüglich stellt Lahno heraus, dass Erwartungen eine Überzeugung über den alles in allem „wahrscheinlichen Verlauf der Welt“[292] darstellen. Wer etwas erwartet, geht davon aus, starke, durch sein Wissen über die Welt gestützte, Gründe für diese Erwartung zu haben. Eine Erwartung sei demnach eine Haltung mit der – zumindest den Anspruch nach – letztlich das Bestehen starker epistemischer Gründe angenommen ist, die dem Subjekt der Erwartung eine entsprechende Sicherheit geben. Vertrauen hingegen sei eine Haltung, die sich idealtypisch in durch epistemische Unsicherheit geprägten Situationen zeige. Insofern man vertraut, drückt dies aus, dass es gerade keine epistemischen Anhaltspunkte für Annahmen über den zukünftigen Weltverlauf gibt: Vertrauen sei damit „insgesamt weniger eine Sache des Wissens und der Überzeugung, als vielmehr eine der Einstellung und des Handelns in einer unüberschaubaren Situation.“[293]

Mit diesem Punkt fängt Lahno sachlich gesehen ein oft festgestelltes begriffliches Merkmal

damit meint, beziehen wir uns kurz auf das Beispiel in der Einleitung: Wenn ich sage, dass ich Peter vertraue, dass er mich wecken wird, dann impliziert dies, dass ich darauf vertraue, dass er mich wecken wird. Der umgekehrte Schluss jedoch ist nicht möglich. Dieser propositionale Aspekt, das Vertrauen-darauf, ist jedoch schlichtweg gleichbedeutend mit dem, was wir oben als Sich-Verlassen bezeichnet haben und beschreibt die allgemeine Struktur vom generischen Vertrauen, unter die auch interpersonelles Vertrauen fällt (vergleiche die Grafik 1). Deswegen werden wir statt Lahnos Begriff des Vertrauen-darauf zu übernehmen, stattdessen vom Vertrauen im generischen Sinne sprechen.

290So übersetzen wir hier Lahnos pragmatistisch orientierte Formulierung: „Er wird sich darauf einrichten, dass das Objekt seiner Erwartung realisiert wird. Genau dies tut derjenige, der auf etwas vertraut ebenfalls“ (Lahno 2002, 157).

291Ebd., 159.

292Ebd., 158.

293Ebd., 158.

von Vertrauen ein. Etwas ist also an dieser Charakterisierung dran. Worauf es uns im Folgenden aber ankommt, ist nachzuweisen, dass der daran anknüpfende Schluss, dass Vertrauen nicht notwendigerweise mit einer Erwartung einhergehen müsse, zu radikal ist. Vielmehr weist Lahnos Argument eigentlich nur auf, dass Vertrauen eine Erwartung *besonderer Art* beinhaltet. Wir zeigen dies, indem wir versuchen genauer zu verstehen, was Lahno mit seinen Hinweisen zu meinen scheint.

Was genau also scheinen seine Abgrenzungskriterien zu sein? Meines Erachtens können bei Lahno diesbezüglich zwei Punkte identifiziert werden. Zum einen verweist er eben darauf, dass Vertrauen *in unsicheren Kontexten* als eine Erwartung situiert sei. In ähnlicher Form wird diese Idee häufiger geäußert. So wird Vertrauen in der Regel als eine epistemische Haltung verstanden, die weniger große Gewissheit ausdrückt, als dies für eine *vollwertige* Erwartung der Fall sei. Anders gewendet kann man auch sagen, dass Vertrauen *schwächere Rechtfertigungsstandards* habe.[294] So bringt es etwa Karen Jones auf den Punkt, die Vertrauen in diesem Sinne auch als eine Form von Hoffnung versteht (und diese in ihrem Fall von einer Prognose abgrenzt): „we can be justified in trusting even when we would not be justified in predicting a favourable action on the part of the one trusted. Our evidence for trusting need not be as great as the evidence required for a corresponding justified prediction. In this respect trusting is more like hoping than like predicting.“[295]

Dieser Punkt scheint sich mir auch daraus ableitbar, dass mit dem Begriff Vertrauen notwendigerweise ein *Risikobewusstsein* impliziert ist, mit dem Begriff der Erwartung als solchem jedoch nicht. Wir erinnern uns an unsere grundlegende Definition von Vertrauen im generischen Sinne: Wenn man sagt, dass man vertraut, impliziert dies, dass man sich eines erwägenswerten Risikos bewusst ist (auch wenn man letztlich optimistisch bleibt). Risiko wiederum setzt sich, wie wir oben gesehen haben aus zwei Komponenten zusammen: Es ist gewissermaßen eine Konjunktion der Unsicherheit mit einem potenziellen Schaden. Vertrauen geht also begrifflich mit

294 Lahno spricht in diesem Zusammenhang allerdings nicht von Rechtfertigungsstandards. In seinen begrifflichen Analysen geht er meistens so vor, dass er sich die Frage vorlegt, zur Beschreibung welcher Situationen man *typischerweise* welche Begriffe *faktisch* verwendet. Wir verstehen begriffliche Analyse hier aber nicht so, dass es darum ginge eine bloße *Korrelation* zwischen der üblichen Verwendung eines Begriffes und seinem Verwendungskontext festzustellen. Vielmehr gehe ich, in Anknüpfung an einen Grundgedanken der normativen Sprachpragmatik, davon aus, dass die Verwendung eines Begriffes an gewisse *normative* Regeln gebunden ist und die begriffliche Klärung, die wir hier vornehmen, eine solche Regel verdeutlicht: Zu sagen, dass man sich sicher sei, dass p der Fall ist, ist dann *gerechtfertigt*, wenn man starke Gründe dafür hat, anzunehmen, dass p der Fall sei. Bei Vertrauen sind diese Rechtfertigungsstandards geringer.

295 Jones 1996, 15. Man kann allerdings wiederum fragen, ob der Begriff der Prognose hier gut gewählt ist. Prognosen können nämlich auch in schwach epistemisch gerechtfertigter Form auftreten, etwa in Form bloßer Hypothesen oder praktischer Annahmen.

einer gewissen epistemischen Unsicherheit einher. Damit steht es im Gegensatz zu epistemischen Haltungen wie Wissen oder einem Sich-sicher-sein. Aus diesem Grund mag es auch irritieren, wenn jemand sagt, er sei sich sicher, dass p passieren wird, obwohl er nur schwache Evidenz dafür hat, die dafür spricht, dass p passieren wird. Wenn jemand hingegen in einer gleichwertigen epistemischen Situation sagt, er vertraue darauf, dass p passieren wird, scheint dies eine nachvollziehbare Redeweise darzustellen. Dies weist darauf hin, dass Vertrauen begrifflich gesehen keinen so hohen Rechtfertgigungsansprüchen untersteht. In dieser Hinsicht begründet sich also ein Unterschied zwischen *Vertrauen* und epistemischen Einstellungen wie *Sich-sicher-sein* oder gar *Wissen.*[296] Ich gehe davon aus, dass es *dieser* Punkt ist auf den Lahno, Jones und andere zurecht hinweisen. Vertrauen ist also eine epistemische Haltung, die von anderen epistemischen Haltungen anhand der Höhe des epistemischen Rechtfertigungsanspruches unterschieden werden kann.

Vertrauen jedoch im selben Sinne von einer Erwartung, Überzeugung oder Prognose abgrenzen zu wollen, wie es Lahno und Jones machen, scheint mir terminologisch verwirrend. Der Begriff der Erwartung etwa ist meinem Eindruck nach in seiner alltagssprachlichen Bedeutung schlichtweg neutral bezüglich seines epistemischen Anspruches, das heißt, mit ihm ist nicht notwendigerweise ein besonders hoher (denken wir an eine Erwartung im Sinne einer Hypothese) noch ein besonders niedriger Standard (denken wir an die Erwartung, dass es schmerzen wird die heiße Herdplatte anzufassen) verbunden. Es scheint mir deshalb von der Wortwahl her naheliegender, die Zusammenhänge vielmehr so zu formulieren: Vertrauen geht mit einer Erwartung einher, auch wenn diese spezifischer qualifiziert werden kann, nämlich als eine Erwartung, die nicht so hohen Rechtfertgigungsstandards unterliegt wie etwa Wissen.

Ähnlich sieht es beim zweiten Abgrenzungskriterium aus, das auch als unmittelbar mit dem ersten zusammenhängend verstanden werden kann. Insofern Vertrauen relativ schwache Rechtfertigungsstandards hat, kennt es auch einen größeren *Rationalitätspielraum*. „Überzeugt sein“ untersteht der Anspruch auf starken Gründen zu beruhen. Vertrauen kann man aber auch

296So kann man etwa bereits schon den Hinweis von Simmel lesen, dass Vertrauen zwischen Wissen und Nichtwissen einzuordnen sei: „Vertrauen, als die Hypothese künftigen Verhaltens, die sicher genug ist, um praktisches Handeln darauf zu begründen, ist als Hypothese ein mittlerer Zustand zwischen Wissen und Nichtwissen um den Menschen. Der völlig Wissende braucht nicht zu vertrauen, der völlig Nichtwissende kann vernünftigerweise nicht einmal vertrauen“ (Simmel 1992, 393). Guy Longworth zieht auf ähnliche Weise eine erhellende Parallele zum Begriff des Glaubens in der Kantischen Erkenntnistheorie. Glauben sei demnach eine Zwischenkategorie zwischen Wissen und Meinen: „one *has faith* that *p* where one takes it to be true that *p* on sufficient subjective grounds, in the consciousness that one’s objective grounds do not furnish a theoretical demonstration of truth.“ (Longworth 2012, 15). Im Glauben gehe man zwar davon aus, dass das Geglaubte der Fall ist – man hat subjektive Gründe und insofern steht es „über“ bloßem Meinen. Zugleich sei Glauben von dem Bewusstsein begleitet keine objektiven Gründe zu haben, die die Wahrheit des Geglaubten beweisen können. Insofern stehe es „unter“ Wissen. Siehe auch die Bemerkungen von: Endreß: 2002, 73f; Faulkner 2007, 897. Diese gehen in ähnlicher Weise davon aus, dass sich Vertrauen durch einen geringeren Grad an Gewissheit auszeichnet.

schon auf der Basis schwächerer Gründe. Deshalb kann der vertrauende Akteur auch eine „aktivere“ Rolle einnehmen: Für Vertrauen gibt es einen größeren rationalen Spielraum als für Überzeugt-Sein. In diesem Sinne interpretiere ich Lahnos Hinweis, dass Vertrauen mehr eine Sache „der Einstellung und des Handelns“ sei. Abgesehen von der Frage jedoch, ob diese Sichtweise plausibel ist, scheint sie mit Blick auf den Versuch, Vertrauen von Erwartungen abzugrenzen irreführend. Einfach ausgedrückt stellt sich auch wenn es bei Vertrauen diesen größeren Spielraum geben mag, die Frage: Den größeren Spielraum für *was – wenn nicht dafür eine bestimmte Erwartung einzunehmen?*[297] Auch an diesem Punkte finde ich es deshalb irreführend zu sagen, dass Vertrauen von einer Erwartung abzugrenzen sei.

Die beiden Punkte zeigen, dass hinter den Überlegungen Lahnos zwar m.E. begrifflich richtige Einsichten liegen. Diese weisen Vertrauen jedoch vielmehr als eine spezifische kognitive Einstellung aus: Vertrauen ist eine Erwartung, die ein Risikobewusstsein impliziert, insofern sie auch eher niedrigen epistemischen Rechtfertigungsstandards unterliegt und damit letztendlich einen größeren rationalen Spielraum zulässt. Es ist damit zurecht von gewissen anderen epistemischen Einstellungen wie einem Sich-sicher-sein oder Wissen abzugrenzen. Es von Erwartungen *gänzlich* abzugrenzen, scheint mir aber terminologisch verwirrend und droht Vertrauen als eine gar nicht mehr kognitive Haltung auszuweisen. Letzteres ist auch die Schlussfolgerung die Lahno letztendlich zieht: „Vertrauen ist folglich *nicht mit logischer Notwendigkeit* an eine Vertrauenserwartung im Sinne einer kognitiven Überzeugung gebunden.“[298] Diese Schlussfolgerung ist auch der entscheidende Grund, weshalb wir die Lahno‘sche Emotionstheorie des Vertrauens in Kapitel 8.1 ablehnen werden.

Es soll abschließend noch einmal betont werden, dass mit diesen Überlegungen nicht gesagt werden soll, dass Vertrauen zu einer bloß epistemischen Einstellung, einer Überzeugung, reduziert werden könnte. Dies soll nicht als Ergebnis dieses Unterkapitels missverstanden werden. Wie wir gerade im Zweiten Teil dieser Arbeit herausgestellt haben, stellt Vertrauen mehr als eine bloß theoretische Überzeugung dar. Wofür wir hier aber argumentiert haben ist, dass Vertrauen notwendigerweise eine Überzeugung oder Erwartung *impliziert*. Vertrauen *umfasst* also auch eine bestimmte epistemische Einstellung.[299]

297In einer späteren Passage schreibt Lahno selber zwar, dass Vertrauen eine Erwartung „induzieren“ könne (ebd.: 210). Er scheint dies jedoch als einen bestimmten kausalen Mechanismus und nicht als einen begrifflichen Zusammenhang aufzufassen. Genauer gehen wir auf dieses Problem unten in der Diskussion von Lahnos Emotionstheorie des Vertrauens ein.

298Ebd., 220; Hervorhebung von mir.

299Ein im Übrigen weiteres, besonders eindrückliches Argument für den Kognitivismus, auf das auch Pamela Hieronymi abhebt, bezieht sich auf den Bereich des Sprechervertrauens. Wenn uns jemand etwas mitteilt, zielt sie darauf ab, dass wir nicht nur so tun als ob wir ihr glauben, sondern dass wir ihr *tatsächlich* glauben. Angenommen

7.3 Epistemische Gründe und der epistemische Anspruch im Vertrauen

Aufbauend auf dem kognitivistischen Verständnis von Vertrauen, für das wir soeben argumentiert haben, widmen wir uns nun der Frage danach, was die *Gründe richtiger Art* für Vertrauen darstellen. Folgenden Argumentationsgang werden wir dafür in diesem Unterkapitel entfalten: Insofern Vertrauen wesentlich die Überzeugung beinhaltet, dass die Vertrauensperson vertrauenswürdig ist, liegt es nahe zu sagen, dass es sich bei den „Gründen richtiger Art“[300] für Vertrauen um solche handelt, die ebendiese Überzeugung als wahr ausweisen. Hierbei handelt es sich demnach um epistemische Gründe. Wenn zudem Vertrauen als eine nicht willkürliche Haltung verstanden werden soll, ergibt sich außerdem, dass Vertrauen zumindest dem Anspruch nach stets auf epistemischen Gründen beruhen muss.

Um dies im Einzelnen etwas näher zu erläutern, widmen wir uns zunächst dem Terminus der Gründe richtiger Art, und wie diese von Gründen falscher Art abzugrenzen sind. Unter Gründen falscher Art versteht man grundlegend solche Gründe, die zwar für eine bestimmte mentale Einstellung, etwa eine bestimmte Überzeugung, sprechen, allerdings ohne, dass man sinnvoll *aus* diesem Grund diese mentale Einstellung gewinnen kann. Ein Beispiel: Nehmen wir etwa an, ein Dämon verspricht mir unermesslichen Reichtum, wenn ich davon überzeugt bin, dass es in Köln freilebende Einhörner gibt. Es sieht nun so aus, als würde mir das Versprechen des Dämons einen Grund für eine bestimmte Überzeugung geben, immerhin spricht nun etwas gewissermaßen dafür, diese Überzeugung zu haben. Allerdings stellt das Angebot einen Grund falscher Art dar: Ich kann nämlich nicht *aus* diesem Grund die Überzeugung haben, dass in Köln Einhörner leben.

Zur Frage, wie genau der Unterschied zwischen Gründen falscher und richtiger Art zu verstehen ist, hat Pamela Hieronymi[301] einen attraktiven Vorschlag gemacht. Er lautet in aller Kürze wie folgt: „the right kind of reasons for (or against) an attitude are *any* that bear (or are taken to

eine Freundin von mir wird eines Verbrechens beschuldigt und teilt mir daraufhin mit, dass sie unschuldig sei. Dann gelte: „Your friend does not want you merely to rely on her claim from a participant stance, to work the supposition that she is innocent into your plans with a readiness to feel betrayed if she is lying—that would be giving her the benefit of the doubt, acting as though she were innocent. Your friend wants you to trust that she *is* innocent; she wants you to believe her.“ (Hieronymi 2008, 219).

300 Die Diskussion um Gründe richtiger und falscher Art (*wrong and right kind of reasons*) hat sich im Ausgang von Scanlons sogenanntem *buck-passing-account* in der Werttheorie entsponnen. Den Ausgangspunkt der Debatte stellt der Artikel *The Strike of the Demon* von Wlodek Rabinowicz's and Toni Rønnow-Rasmussen (2004) dar. In jüngerer Zeit werden Gründe richtiger und falscher Art aber auch in anderen philosophischen Kontexten diskutiert, etwa in Bezug auf Überzeugungen. Eine Übersicht über die Debatten findet sich auch in: Gertken/Kiesewetter 2017.

301 Hieronymi 2005; Hieronymi 2013.

bear) on a question (or set of questions) the positive settling of which amounts to forming the attitude.“[302] Die Idee ist dabei zunächst, dass verschiedene Arten von Einstellungen (Bedürfnisse, Überzeugungen, Intentionen etc.) als Festlegungen in Bezug auf je spezifische Fragen individuiert werden können. Demnach kann beispielsweise eine Überzeugung, dass p der Fall ist, als positive Festlegung in Bezug auf die Frage: „Ist es wahr, dass p?“ verstanden werden. Eine Absicht ϕ zu tun kann als positive Festlegung in Bezug auf die Frage „Soll ich ϕ tun?“ verstanden werden. Gründe richtiger Art können in diesem Zusammenhang als solche Gründe verstanden werden, die *für oder gegen eine bestimmte Antwort auf eine solche Frage sprechen* – in Bezug auf Überzeugungen beispielsweise sind dies etwa Gründe, die auf Indizien rekurrieren, die dafür sprechen, dass p wahr ist. Gründe falscher Art weisen demgegenüber *nicht* auf eine Antwort einer solchen Frage hin. Vielmehr weisen sie direkt eine bestimmte Einstellung *als solche* als wünschenswert aus. In Bezug etwa auf eine Überzeugung, dass p der Fall ist, könnten Gründe falscher Art solche sein, die herausstellen, dass es *angenehm* wäre, von p überzeugt zu sein. Gründe falscher Art weisen also eine bestimmte Einstellung *als solche* direkt als wünschenswert oder nicht-wünschenswert aus, ohne dabei auf die sie konstituierende Frage zu antworten. Sie sehen damit gewissermaßen wie Gründe aus, immerhin sprechen sie für etwas, etwa dafür die Überzeugung zu haben, dass p der Fall ist. Es ist jedoch nicht möglich unter Bezugnahme auf diesen Grund zur Überzeugung zu *kommen*, dass p der Fall ist.

Vor dem Hintergrund dieser Klärung können wir uns nun fragen, was die Gründe richtiger Art für Vertrauen darstellen. Betrachten wir dazu zunächst den Fall nicht-personalen Vertrauens. Nicht-personales Vertrauens, etwa das Vertrauen, dass x passieren wird, geht, so hatten wir oben gezeigt, mit der Überzeugung einher, dass x passieren wird. Damit erweisen sich als dessen Gründe richtiger Art solche Gründe, die auf die Frage, ob x passieren wird, eine positive Antwort nahelegen. Die richtigen Gründe sind hier also solche, die dafür sprechen, dass es *wahr* ist, dass x passieren wird. Für den Fall interpersonellen Vertrauens gilt zudem: Insofern dieses die Überzeugung beinhaltet, dass die Vertrauensperson vertrauenswürdig ist, stellt es wesentlich eine positive Beantwortung der Frage dar „Ist die Vertrauensperson vertrauenswürdig?“. Gründe richtiger Art für interpersonelles Vertrauen sind demnach solche, die die andere Person als vertrauenswürdig ausweisen. Sowohl die Gründe nicht-personalen als auch interpersonellen Vertrauens fallen damit unter die Klasse von Gründen, die ich *epistemische Gründe* nenne. Damit sind schlichtweg solche Gründe gemeint, die *dafür oder dagegen sprechen, dass etwas Bestimmtes*

302Hieronymi 2013, 118.

der Fall ist.[303]

Im Fall interpersonellen Vertrauens können solche epistemischen Gründe, die für die Vertrauenswürdigkeit einer Person sprechen, im Einzelnen äußerst vielfältig ausfallen: Diese können etwas unmittelbar an der Vertrauensperson Wahrnehmbares sein (ein sozialer Marker, ein Kleidungsstück, das ihre Rolle kennzeichnet, eine offene Art des Auftretens).[304] Ein solcher Grund kann auch das vergangene Betragen einer Person darstellen, etwa die Tatsache, dass sich diese bisher immer als vertrauenswürdig erwiesen hat. Ebenso kann auch ein Element des Kontextes ein epistemischer Grund sein, der darauf schließen lässt, dass die Person vor mir vertrauenswürdig ist. So mag etwa der Umstand, dass man in einer Gesellschaft mit hohem Vertrauensklima und funktionierenden Institutionen lebt, auch für die Vertrauenswürdigkeit der Person vor mir sprechen.[305] Epistemische Gründe sind hier wie gesagt recht weit verstanden als Gründe, die für oder gegen die Vertrauenswürdigkeit des Interaktionspartners sprechen. Auf welche Art von Evidenz genau sich ein solcher Grund beziehen mag, lassen wir damit an dieser Stelle also offen.[306]

Gegen die These, dass epistemische Gründe die Gründe richtiger Art für Vertrauen darstellen, mag es Einwände geben, von denen wir an dieser Stelle zwei adressieren wollen. Beide beziehen sich speziell auf den Fall interpersonellen Vertrauens. So mag man sich erstens fragen, ob es nicht neben diesen epistemischen Gründen, die für die Vertrauenswürdigkeit der anderen Person sprechen, *auch andere Gründe* für Vertrauen geben mag. Dies mag man etwa aus folgender Überlegung heraus vermuten: So scheint es plausibel zu sagen, man vertraue einer Person, da dieses Vertrauen etwa nützlich sei, oder einen funktionalen oder intrinsischen Wert habe o.Ä. Solch eine Art von Begründung weist damit allgemein darauf hin, was überhaupt der Punkt am Vertrauen

303 Diese Beschreibung einer Überzeugung steht auch im Zusammenhang mit einer berühmten von David Velleman vorgenommenen Charakterisierung von Überzeugungen: „believing a proposition entails regarding it as something that one is right to regard as true“ (1992, 15). Eine solche Berechtigung, etwas als wahr anzusehen, bezieht sich aber wesentlich auf Gründe, die für die Wahrheit des Geglaubten p sprechen. Und insofern sind es Gründe, die für eine positive Antwort auf die Frage „Ist es wahr, dass p?“ sprechen.

304 Die philosophische Literatur kapriziert sich zugegebenermaßen auf diese Arten von Gründen. Trudy Govier listet demgegenüber 5 verschiedener Arten epistemischer Gründe für Vertrauen in eine Person P auf. Diese beruhen entweder: auf direkter personaler Kenntnis von P (1), auf dem Wissen anderer über P (2), auf Buchwissen über P (3), auf Wissen durch andere Medien über P (4), oder auf der sozialen Rolle von P (5) (Govier 1998, Kap. 7.). Eine in meinen Augen wichtige Art epistemischer Gründe ist damit jedoch ausgelassen, und zwar solche, die sich auf den sozialen Kontext der Begegnung mit der anderen Person beziehen (etwa auf das allgemeine Vertrauensklima).

305 Diese Arten von Gründen sind insbesondere für Formen sogenannten generalisierten oder „versachlichten Vertrauens“ (Simmel) entscheidend. Damit ist Vertrauen gemeint, dass sich auf Personen in ihren Professionen oder gesellschaftlichen Funktionen bezieht und keiner individuenbezogenen, personaler Kenntnisse bedarf (vgl. Simmel 1992, 394). Diese Gründe beziehen sich vielmehr auf die Vertrauenswürdigkeit eines generischen Personentypus (etwa bestimmter Experten). Zur Rolle von Institutionen zur Begründung der Annahme, dass Experten vertrauenswürdig sind, siehe auch: John 2018, 78.

306 Neben diesen epistemischen Gründen als solchen ist natürlich auch deren Kohärenz untereinander relevant, um Vertrauen gut zu begründen.

ist. Diese Begründung charakterisiert den *Sinn* von Vertrauen, indem sie auf instrumentelle oder intrinsische Zwecke von Vertrauen rekurriert. Der Punkt ist aber, dass diese „Gründe“ nicht die *richtige* Grundlage darstellen können, auf der sich das Vertrauen in eine andere Person *konstituiert*. Ich mag den Sinn sehen, den es hätte, Vertrauen in eine bestimmte andere Person zu haben. Aber dies würde keinen Grund darstellen, *aus dem* ich Vertrauen in die andere Person haben kann. Vertrauen in eine andere Person zu haben, hieße nämlich, die Überzeugung zu haben, dass die konkrete andere Person auch vertrauenswürdig *ist*. Dafür aber braucht es epistemische Gründe, die die andere Person als vertrauenswürdig ausweisen.[307] [308] Vielleicht möchte man hier insistieren, und darauf hinweisen, dass man doch problemlos sagen könne: „Die Beziehung zu ihm ist mir wichtig, aus diesem Grund vertraue ich ihm.“ Darauf ist zu entgegen, dass hier die Ambiguität der Alltagssprache ein falsches Bild nahelegt: Unter „Vertrauen“ kann in dieser Aussage nämlich allenfalls ein Vertrauensakt (oder in Pamela Hieronymis Worten bloßes *entrusting*) gemeint sein. Aufgrund der Relevanz der Beziehung als solcher, kann man nicht zu einer vertrauenden Haltung gelangen, man kann nicht *aus* diesem Grund Vertrauen in die andere Person haben. Man kann höchstens so tun, als ob man vertraute.

Ein zweiter Einwand gegen die These der epistemischen Gründe als Gründe richtiger Art mag sich auf unsere Anerkennungstheorie im zweiten Teil dieser Arbeit beziehen. Dort hatten wir ausgeführt, dass Vertrauen in der Anerkennung einer normativen Erwartung der Vertrauensperson besteht. Damit stellt sich die Frage, wie dies mit der hier vertretenen Ansicht vereinbar sei, dass Vertrauen wesentlich auf epistemischen Gründen beruhe. Auf diesen Themenkomplex gehen wir im vierten Teil dieser Arbeit noch genauer ein. An dieser Stelle sei aber bereits angemerkt, dass ich

307 Oben (Exkurs 4.3) haben wir allgemein Gründe mit Anscombe als „Erwünschbarkeitscharakterisierungen“ definiert: Gründe weisen auf, inwiefern bestimmte Handlungen oder Haltung im Lichte eines bestimmten Gutes erwünschbar sind. In diesem Zusammenhang hatten wir auch schon darauf hingewiesen, dass an der Formulierung der Gründe *nicht explizit ersichtlich sein muss*, dass diese sich an einem bestimmten Wert orientieren. Gründe mögen demnach zwar immer noch wesentlich einen Bezug zu einem Wert haben, aber: sie können sich auch *indirekt* an diesem orientieren. Dies scheint mir grundlegend für epistemische Gründe als Gründe richtiger Art zu gelten. Solche Gründe verweisen lediglich auf Evidenzen, die dafür sprechen, dass etwas der Fall ist („Aus diesem Indiz lässt sich darauf schließen, dass p zutrifft“). Diese beziehen sich nicht *ausdrücklich* auf irgendein Gut. Dieses steht in der Urteilsbildung durch epistemische Gründe aber im Hintergrund. So orientiert sich etwa auch epistemische Urteilsbildung konstitutiv am Wert der Wahrheit, ohne dass epistemische Gründe direkt und ausdrücklich auf dieses Gut Bezug nehmen. Siehe hierzu auch die Überlegungen von Anselm Müller zum Begriff der „geistigen Teleologie“ (Müller 1992).

308 Wir haben in diesem Argumentationsgang die epistemischen Gründe, die für die Vertrauenswürdigkeit der anderen Person sprechen als die Gründe richtiger Art für Vertrauen ausgewiesen, indem wir allein von einer begrifflichen Analyse einer vertrauenden Einstellung ausgegangen sind. Der Gedanke lässt sich aber auch von einer anderen, wenn man so will „funktionalen“ Richtung kommend stützen. Demnach hängen am Vertrauen, wie oben hinlänglich gezeigt, verschiedene Arten von Werten. Diese lassen sich aber überhaupt nur realisieren, wenn das Vertrauen aber auch *gelingt*. Die entscheidende Gelingensbedingung für Vertrauen ist aber, ob die andere Person vertrauenswürdig ist. Auch mit Blick auf zu realisierende Güter, sind also auch solche Gründe, die für die Vertrauenswürdigkeit der anderen Person sprechen, die entscheidenden.

hinter diesem Einwand eine Verwechslung vermute: Oben haben wir vielmehr aufgezeigt, dass Vertrauen in der Anerkennung einer normativen Erwartung einer Person *besteht*. Etwas anderes wäre es zu behaupten, dass die normative Erwartung der anderen Person einen *Grund* darstellt, diese zu akzeptieren. Diese Annahme würde in der Tat wiederum einen Grund falscher Art präsentieren. Ich kann nämlich nicht einfach jemandem deshalb vertrauen, da er es von mir erwartet.[309]

Wir können nun einen Argumentationsschritt weiter gehen und zeigen, dass Vertrauen zudem stets dem *Anspruch* nach ausreichend durch die genannten epistemischen Gründe gerechtfertigt ist. Es ist nicht so, dass Vertrauen nun entweder durch solche epistemischen Gründe gestützt sein kann, es *genauso gut* aber auch nicht sein könnte. Dies ergibt sich daraus, dass wir Vertrauen nicht einfach als eine Haltung verstehen, zu die der Vertrauende radikal willkürlich kommen kann. Vertrauen kann kein bloß a-rationales Phänomen darstellen. Vertrauen verstehen wir vielmehr als zumindest dem *Anspruch* des Vertrauenden nach nicht vollkommen arbiträr. Wir verstehen Vertrauen so, dass aus Sicht eines Vertrauenden etwas dafür sprechen muss, dass sein Vertrauen angemessen ist, sprich, dass es Gründe für dieses gibt. Wenn eine Person einer anderen Person A, aber nicht der Person B vertraut, dann versteht sie dies zumindest dem Anspruch nach nicht als bloß willkürliches, sondern als situativ begründetes Verhalten.

Insofern nun aber nachgewiesenermaßen epistemische Gründe die Gründe richtiger Art für Vertrauen darstellen, ergibt sich insgesamt, dass Vertrauende zumindest dem Anspruch nach ihr Vertrauen als epistemisch begründet verstehen. Es zeigt sich damit, dass Vertrauen inhärent mit einem Anspruch auf epistemische Begründetheit einhergeht. Damit ergibt sich also insgesamt folgende These in Bezug auf die epistemologische Struktur von Vertrauen:

THESE DES INNEREN EPISTEMISCHEN ANSPRUCHS VON VERTRAUEN

Vertrauen ist mit einem inhärenten Anspruch unterlegt epistemisch begründet zu sein.

Der Punkt, dass Vertrauen mit einem inneren Anspruch auf situativer Angemessenheit einhergeht, stellt genau genommen keine Besonderheit des Vertrauensbegriffes als solchem dar,

309 Vergleiche ähnliche Argumente, inwiefern praktisch-normative Anforderungen also solche, zumindest nicht unmittelbar, Gründe liefern können, aus denen man eine bestimmte Überzeugung hat (Vgl. Zagzebski 2012, 99-102).

sondern sagt vielmehr etwas ganz Grundlegendes über die Grammatik menschlichen Urteilen[310] oder menschlichen Handeln überhaupt aus. Julian Nida-Rümelin bringt diesen Gedanken etwa in Bezug auf menschliches Handeln pointiert zum Ausdruck, wenn er schreibt: „Jede Handlung ist Ausdruck einer normativen Stellungnahme der Art: *Ich bin überzeugt, dass diese Handlung richtig ist"*.[311] Er drückt damit aus, dass menschliches Handeln stets konstitutiv mit einem Anspruch unterlegt ist, der Situation angemessen zu sein. So erklärt sich auch, weshalb Anscombe die Anwendbarkeit der Warum-Frage als konstitutiv für absichtliche menschliche Handlungen ansieht[312]: Für Anscombe grenzen sich Handlungen etwa von bloßen Naturgeschehen dadurch ab, dass sich an den Handelnden sinnvollerweise die Frage richten lässt „Warum hast du ϕ getan?". Diese Frage lässt sich so verstehen, dass sie den impliziten Anspruch der Handelnden darauf aufgreift, Gründe für ihre Handlungen zu haben, und nach dessen Einlösung verlangt. Man könnte deshalb in Analogie zu Anscombes Handlungstheorie sagen, dass auch für Vertrauen wesentlich ist, dass sich eine Frage wie „Warum vertraust du P?" sinnvoll an den Vertrauenden stellen lässt. Und diese Frage gestaltet sich hier als die Frage nach den epistemischen Gründen, die dafür sprechen, dass der Vertrauenspartner vertrauenswürdig ist.

Diese Überlegung soll nicht verneinen, dass Vertrauen auch an diesen Ansprüchen gemessen scheitern und sich also als irrational erweisen kann. Die Idee ist hier vielmehr, dass Vertrauen mit einer intrinsischen Orientierung an der epistemischen Begründetheit einhergeht. Es kann sich deshalb als irrational erweisen, ist damit aber gerade nicht einfach nur *a*-rational. Metaphorisch gesprochen, ist Vertrauen stets in einem „Raum epistemischer Gründe" situiert. Die Vertrauende kann dabei dem internen epistemischen Anspruch nach Begründetheit nicht gerecht werden, ihr Vertrauen sich deshalb auch als naiv oder unvorsichtig herausstellen. Mehrere Weisen dem Anspruch nicht gerecht zu werden sind hier denkbar: Es kann sein, dass der Vertrauenden aufgeht, dass es überhaupt keinen guten Gründe für ihr Vertrauen gibt, und sie dieses besser aufgeben sollte. Ebenso ist denkbar, dass sie Gründe nennen kann, diese sich jedoch als äußerst schwach oder vielleicht sogar nichtig herausstellen. Die Gründe, die die Vertrauende annahm, können ebenso in Zweifel gezogen, unterminiert oder mit Gegengründen übertrumpft werden. In all diesen Fällen mag die Vertrauende zwar an ihrem *Anspruch* scheitern und ihr Vertrauen für die

310Siehe diese Erläuterung etwa zum Begriff des Urteilens als einem sogenannten normativen Akt von Andrea Kern: „Ein Akt, zu dessen Vollzug es gehört, daß das Subjekt sich dabei einem bestimmten Maßstab unterwirft, ist ein normativer Akt […] Ein Akt, zu dessen Vollzug es überdies gehört, daß das Subjekt in seinem Vollzug von diesem Maßstab geleitet wird, ist nicht nur einfach normativ aus der Perspektive von irgendjemand, sondern er ist normativ aus der Perspektive des Subjekts dieses Aktes selbst" (Kern 2006, 32).

311Nida-Rümelin 2009, 105.

312Anscombe 2000.

Zukunft womöglich überdenken. Dies heißt aber gerade nicht, dass sie damit „eigentlich“ gar nicht oder „nicht wirklich“ vertraut hätte.[313]

An diesem inneren Anspruch gemessen, lässt sich im Übrigen das Vertrauen einer Person auch kritisieren. Das können wir etwa tun, wenn wir das Gefühl haben, dass sich eine Person nicht bemüht diesem Anspruch auf epistemischer Begründetheit des Vertrauens gerecht zu werden. Wir werfen einer Person dann etwa vor, dass sich diese unbedacht oder naiv verhalte etc. Auch an solchen Reaktionsweisen lässt sich ablesen, dass Vertrauen mit einem internen Anspruch darauf einhergeht auf epistemischen Gründen zu beruhen.

Vielleicht mag ein Einwand an dieser Stelle lauten, dass man doch auch vollkommen rational vertrauen könne, ohne überhaupt irgendwelche Gründe dafür zu haben. So mag man einwenden, dass Vertrauen doch auch insoweit rational sein kann, als *keine negativen* epistemischen Gründe vorliegen, also keine Gründe die *gegen* die Vertrauenswürdigkeit des anderen sprechen. Mit anderen Worten: Vertrauen könne ja auch soweit rational sein, als keine eindrücklichen Gründe für Zweifel vorliegen. Gemäß dieser Vorstellung wäre Vertrauen eine Art *prima facie* gerechtfertigte *default*-Einstellung, die nur insofern einzuklammern und zu überdenken sei, als sich konkrete Anlässe für Zweifel ergeben. Diese Konzeption der epistemischen Konfiguration einer vertrauenden Einstellung können wir auch als *kontextualistische Konzeption*[314] bezeichnen. Verschiedene Versionen derselben findet man insbesondere in der Zeugenschaftsdebatte unter den nicht-reduktionistischen Sichtweisen[315], als auch im Vertrauensdiskurs, wenn auf die fundierende, welterschließende Rolle von Vertrauen etwa in Form von Grundvertrauen hingewiesen wird.[316]

Beide Versionen werden wir im weiteren Verlauf der Arbeit aufgreifen und genauer unter die Lupe nehmen. Tatsächlich wird sich zeigen, dass die kontextualistische Konzeption, insbesondere in bestimmten Vertrauenskontexten, eine plausible Beschreibung der epistemischen

313 Dieser Punkt spiegelt eine häufige Fehlinterpretation der Anscombe‘schen Handlungstheorie (vgl. Müller 2015, 263f.). Auch Anscombe selber macht klar, dass nur eine *Zurückweisung* der Warum-Frage anzeigt, dass gar keine absichtliche Handlung vorlag. Bereits der Versuch hingegen die Frage zu beantworten, auch wenn dies nicht wirklich gelingen sollte, akzeptiert die Frage *als kategorial angemessen*, und damit auch, dass eine absichtliche Handlung vorlag.

314 Diesen Begriff verwenden wir in Anlehnung an die Position des Kontextualismus, die als Antwort auf den Skeptizismus in der Erkenntnistheorie vertreten wird. Sie wird vorrangig mit der antiskeptischen Strategie verbunden wird, die John Austin in seinem Aufsatz *Other Minds* (Austin 2000) vertritt. In Bezug auf die Herausforderung des Skeptizismus zu überwinden, mag diese Position übrigens kritikabel sein (vgl. Kern 2006, 112-130). Das hindert einen jedoch nicht daran, die grundlegende kontextualistische Idee zur Beschreibung der epistemischen Struktur von Vertrauen heranzuziehen.

315 Dies ist insbesondere für die sogenannte *entitlement*-Sichtweise der Fall, auf die wir unten im vierten Teil genauer eingehen werden. Die seltene Ausnahme einer nicht-reduktionistischen Position, die zudem postuliert, dass Vertrauen positive Gründe braucht, vertritt Paul Faulkner (Faulkner 2011).

316 Siehe etwa: Hertzberg 1988. Vgl. Endreß 2013.

Struktur von Vertrauen bereitstellt. Dennoch soll schon an dieser Stelle betont sein, dass sie in jedem Fall nicht in Opposition zu der hier vertretenen These steht, dass Vertrauen dem Anspruch nach auf epistemischen Gründen beruht. Diese besagt keinesfalls, dass Vertrauen nur dann rational sei, sofern es sich auf *positive* Gründe stützt, die für die Vertrauenswürdigkeit der anderen Person sprechen. Sie soll vielmehr durchaus zulassen, dass Vertrauen auch dadurch gerechtfertigt sein kann, dass keine bedeutenden negative Gründe, dass keine Gründe für Zweifel vorliegen. Schließelich haben wir lediglich festgehalten, dass, wenn Vertrauen nicht radikal willkürlich oder a-rational verstanden werden soll, es in einer grammatischen Beziehung zu Gründen stehen muss. Die kontextualistische Fassung ist eine Version, diese Beziehung zu Gründen zu verstehen. Dies ist eine besondere, und von der These des inneren epistemischen Anspruches keinesfalls ausgeschlossene, Form unter Bezugnahme auf der Idee von Gründen (hier negativen Gründen), Vertrauen als nicht-willkürlich verständlich zu machen.

7.4 Vertrauen als festgelegte Haltung

Die vorangegangenen Überlegungen haben gezeigt, dass Vertrauen mit der Überzeugung einhergeht, dass die andere Person vertrauenswürdig ist. Diese Überzeugung ist mit einem Anspruch darauf unterlegt, epistemische Gründe für die Vertrauenswürdigkeit der Vertrauensperson zu haben. Bei diesen epistemischen Gründen handelt es sich um die entscheidenden Gründe richtiger Art für Vertrauen. Das genaue *Verhältnis* dieser epistemischen Gründe zum Vertrauen bleibt damit aber noch weitgehend unterbeleuchtet. Auf einen trivialen begrifflichen Punkt in Bezug auf dieses Verhältnis wollen wir in diesem Unterkapitel eingehen. Vertrauen stellt demnach eine Haltung dar, mit der die Frage danach, ob die Vertrauensperson vertrauenswürdig ist, bereits positiv *beantwortet* ist. Insofern sie aber positiv beantwortet ist, ist sie aber trivialerweise mit dem *Erwägen* von Gründen für oder gegen die Vertrauenswürdigkeit unvereinbar. [317]

Wohlgemerkt ist dieser Punkt eine triviale begriffliche Bemerkung: Vertrauen und Zweifel bzw. das Suchen nach Gründen stehen im logischen Widerspruch zueinander. Damit ist aber nicht ausgeschlossen, dass man beispielsweise die Vertrauenswürdigkeit einer Person, der man im Allgemeinen vertraut, hin und wieder hinterfragen kann. Der Punkt läuft also nicht darauf hinaus,

317Im Übrigen zeigt sich an dem bisher Gesagten eine weitere artikulatorische Beschränkung der Rational Choice-Theorie des Vertrauens. Insbesondere kann sie auf der Grundlage ihres Vokabulars nicht den Aspekt beschreiben, dass aus Sicht des Vertrauensgebers, *insofern er vertraut*, Zweifel ausgeblendet sind.

dass man dazu verdammt sei nur „blind“ vertrauen zu können. Er besagt lediglich, dass *sofern* man sein Vertrauen hinterfragt, man gerade nicht (mehr) vertraut. Insofern man zweifelt, klammert man das Vertrauen gewissermaßen ein. Dies ist zumindest der Fall, sofern man es nicht bloß mit rein hypothetischen Zweifeln, sondern vielmehr mit *ernsthaft hinterfragenden* Zweifeln zu tun hat, die man mit Charles Sanders Peirce etwa als „real and living doubt“ bezeichnen könnte.[318] In diesem Punkt unterscheidet sich Vertrauen im Übrigen gar nicht von verwandten epistemischen mentalen Zuständen, etwa wie eine Erwartung oder Überzeugung. Auch diese bestehen darin, dass man sich in Bezug auf eine epistemische Frage („Ist p der Fall?“) *festgelegt* hat, und sind insofern als solche unmittelbar mit dem ernsthaften Erwägen dieser Frage selber inkompatibel.

Es ist wichtig diesen Punkt hervorzuheben, um ihn klarer von dem Punkt abzugrenzen, den wir im folgenden Kapitel diskutieren werden. Man beachte nämlich, dass mit dem hier Gesagten nicht logisch ausgeschlossen ist, dass man zuweilen auch auf der Grundlage von ausdrücklichen Überlegungen dazu gelangen kann, einer Person zu vertrauen. Demgegenüber wurde in der Vertrauensliteratur jedoch mehrfach betont, dass es Situationen gibt, in denen die Tatsache, dass man sich *auf Grundlage expliziter Deliberation* darauf festlegt zu vertrauen, bereits Vertrauen zuwiderläuft oder dieses untergräbt. Zum Vertrauen gehöre demnach zuweilen auch, dass Zweifel nicht nur letztlich *überwunden*, sondern dass sie *gar nicht erst* in Erwägung gezogen werden. Wie sich diese Idee begründet und wie genau sie zu verstehen ist, soll im folgenden Kapitel geklärt werden.

8. Vertrauen und Reflexion

Wie im vorangegangenen Kapitel gezeigt wurde, weist Vertrauen einen analytischen Zusammenhang zu epistemischen Gründen auf. Den epistemischen Aspekt von Vertrauen zu verstehen, bedeutet auch zu verstehen, inwiefern Vertrauen dem Anspruch nach auf solchen Gründen beruht. Dies bedeutet freilich jedoch nicht, dass man nur dann vertraut, wenn man sich seiner Gründe dafür ausdrücklich bewusst ist. Noch weniger heißt es, dass Vertrauen notwendigerweise im Modus ausdrücklicher Überlegungen und Entscheidungsprozesse vollzogen wird. *De facto* manifestiert sich zwar ein Teil unseres Vertrauens in dieser Form. So scheint es, dass wir in bestimmten Kontexten zu manchen Personen erst nach einigem Überlegen und

318Peirce 1992, 115.

Abwägen Vertrauen fassen können. Insbesondere in Vertrauenssituationen, die uns nicht vertraut sind und in denen wir mit unserem Vertrauen ein hohes Risiko eingehen, scheint dies geradezu der erwartbare Normalfall zu sein. Ein Großteil unseres Vertrauens jedoch vollzieht sich in einem präreflexiven Modus. Zumeist wägen wir nicht nur keine Gründe für oder gegen unser Vertrauen ab. Vielmehr noch kommt es uns in der Regel nicht einmal zu Bewusstsein, dass wir überhaupt vertrauen. In vielen alltäglichen Kontexten „vertrauen wir *einfach"*. Um diesen spezifischen *Modus* des unthematischen Vertrauensvollzuges soll es in diesem Kapitel gehen. Im Fokus steht dabei die Frage, wie für diesen speziellen Vertrauensmodus der begriffliche Bezug zwischen Vertrauen und epistemischen Gründen zu spezifizieren ist. Dass Vertrauende zum Teil in spontaner und unreflektierter Weise vertrauen, kann nämlich nicht heißen, dass ihr Vertrauen schlichtweg ein *a-rationales* Phänomen darstellt.

Wie ist es also zu denken, dass Vertrauen unreflektiert aber nicht a-rational vollzogen wird? Diesbezüglich werden wir in diesem Kapitel drei mögliche Modellierungen unreflexiven Vertrauens diskutieren. Dies ist zum einen die Emotionstheorie des Vertrauens, wie sie etwa von Bernd Lahno vertreten wird (8.1). Diese repräsentiert Vertrauen zwar als eine spontane Einstellung. Sie kann jedoch, wie wir zeigen werden, nicht verständlich machen, inwiefern Vertrauen damit nicht einfach eine a-rationale, vollkommen willkürliche Haltung darstellt, die der Vertrauende nicht einmal unter dem Anspruch versteht der Situation angemessen zu sein. Sie droht damit Vertrauen auch nicht mehr als kritikabel und lernfähig ausweisen zu können. Besser geeignet scheint dafür hingegen die Tugendtheorie des Vertrauens, wie sie etwa von Elizabeth Fricker entwickelt wurde (8.2). Diese erfasst den kognitiven Aspekt von Vertrauen und liefert einen geeigneten Ausgangspunktm um nicht-reflexive Vertrauensformen zu konzeptionalisieren. Unterbelichtet bleiben dabei jedoch die genauen Mechanismen des praktischen Lernprozesses dahinter und insbesondere die Rolle die vertrauende oder misstrauende Subjekte in diesem Zusammenhang haben. Damit bleibt etwa der Umstand dunkel, dass wir Akteure auch für ihre präreflexiven Vertrauenseinstellungen verantwortlich machen können.

Wir werden aufgrund dessen vorschlagen, unreflektierte Vertrauens- und Misstrauensformen als *habitualisierte Praktiken* zu fassen (8.3). In Form dieser Kategorie betrachtet, lässt sich der Prozess des Erwerbs unreflektierter Vertrauensformen angemessen beschreiben. Insbesondere wird die Verantwortung von Subjekten im Vollzuge solcher Vertrauenspraktiken intelligibel. Mit der Idee, dass Vertrauen als Praxis habitualisiert werden kann, lässt sich ein Bild von Vertrauen als einem dynamischen und Lernprozessen unterliegenden

praktischem Verhalten von Akteurinnen zeichnen.

In einem weiteren Schritt werden wir zeigen, dass unser alltagssprachlich verankertes Vertrauensverständnis im Kern tatsächlich auf Vertrauen in seinem habitualisierten Modus abzielt (8.4). Vor diesem Hintergrund lässt sich auch klarer nachvollziehen, warum sich in der Literatur verstreut der Gedanke formuliert findet, dass ausdrückliche und ernsthafte Reflexion darüber, ob ein Vertrauenspartner vertrauenswürdig ist, eigentlich mit veritablem Vertrauen unvereinbar sei. Mehr noch, zuweilen wird dies auch mit der Vorstellung verbunden, dass die Explizierung von Vertrauen überhaupt, dieses potenziell unterminieren könne. Oli Lagerspetz etwa, dessen Werk diesbezüglich sicherlich den exponiertesten Referenzpunkt darstellt, hält fest: „a certain lack of reflexion and explicitness is *constitutive* of [...] trustfulness.“[319]. Ähnlich betont Arnon Keren in negativer Formulierung: „trust is undermined by excessive reflection.“[320] Und Martin Hartman hält fest: „Nachdenken so scheint es zerstört Vertrauen, zersetzt es [...]“.[321] Nennen wir die verbindende Hypothese, die sich in diesen Aussagen ausdrückt die *Hypothese der Reflexionslosigkeit* von Vertrauen. Der Inbegriff von Vertrauen würde sich dieser Hypothese folgend in einem Vertrauensverhalten ohne Zögern und Überlegen offenbaren. In seiner wesentlichen Form sei Vertrauen also als eine Art *nicht überlegungsgestützte, spontane Einstellung* zu begreifen. Um die Bedeutung der Hypothese nachzuvollziehen, ist es wichtig, dass hier auf Vertrauen nicht als einer sich in Einzelsituationen und episodisch zeigenden Einstellung referiert wird. Vielmehr wird auf Vertrauen als einer zeit- und kontextübergreifenden festen und habitualisierten Einstellung oder Praxis Bezug genommen.

8.1 Die Emotionstheorie des Vertrauens

Welche konzeptionellen Möglichkeiten gibt es nun also die Form unreflexiven Vertrauens zu fassen? In der philosophischen Vertrauensliteratur gibt es dazu einige einflussreiche Ansätze, die sich als Emotionstherien des Vertrauens klassifizieren lassen. Zu ihren zwei bekanntesten Vertreter:innen gehören Bernd Lahno und Karen Jones.[322] Sie eint die Sichtweise, dass Vertrauen

319Lagerspetz 1998, 21.

320Keren 2014, 14.

321Hartman 2020, 11. In eine ähnliche Richtung weisen auch Bemerkungen in: McGeer (2008, 240); Townley (2011, 23–43); Hawley 2012, 3; Hieronymi (2008, 217). Lahno 2002, 208. Martin Endreß formuliert „dass Vertrauen, sobald man es thematisiert, aus diesem Grunde aufhört, bloßes Vertrauen zu sein“ (Endreß 2001, Fn. 43).

322Lahno (2001, 2002), Jones (1996), siehe auch: Miller (2000). Wir stützen uns hier vorrangig auf den Ansatz von Bernd Lahno, wie er ihn in Lahno (2002) eingeführt hat. An anderer Stelle scheint er die hier formulierte Kritik zu

wesentlich eine Emotion[323] darstelle, und damit insgesamt weniger eine Sache einer kalkulierten Entscheidung auf Grundlage epistemischer Gründe, sondern vielmehr eine Art darstelle seine Umgebung bzw. seinen Vertrauenspartner, in einem bestimmten Lichte *wahrzunehmen*.

Im Folgenden werden wir den Ansatz von Bernd Lahno diskutieren, da dieser am ausführlichsten ausgearbeitet ist. Dieser bezieht sich auf eine aristotelisch angelehnten Theorie der Gefühle[324], nach der diese wesentlich eine wahrnehmungsfärbende oder wahrnehmungsanleitende Dimension haben. Sie konstituieren den Blick, den ein Subjekt auf die es umgebende Situation einnimmt: Welche Aspekte der Umwelt für sie salient und hervorgehoben sind, wie sie diese interpretiert und auch welche Assoziationen und Empfindungen diese für sie hervorrufen. Lahno erläutert diesen Punkt anhand des folgenden Beispieles: Angenommen eine Person hat Angst vor wilden Bären, dann hat diese emotionale Disposition einen unmittelbaren Einfluss darauf, wie die Person einen nächtlichen Wald wahrnimmt und worauf ihre Aufmerksamkeit gelenkt sein wird. Ihre Angst bestimmt worauf sie achtet (mögliche Verstecke von Bären), wie sie bestimmte Aspekte interpretiert (abgebrochene Zweige als mögliche Indizien für Bären) und zu welchen Überlegungen ihre Umgebung sie anregt (Kann ich diesen Baum erklettern?). Gefühle sind diesem Verständnis nach mehr als bloß subjektive Empfindungszustände. Sie sind vielmehr Teil der unmittelbaren und unhintergehbaren *Perspektive*, aus der eine Person ihre Umgebung wahrnimmt. Entsprechend hält Lahno fest: „Gefühle sind wie eine Brille, durch die wir die Wirklichkeit betrachten, ein Rahmen für unsere Aufnahme und Auseinandersetzung mit der Welt.“[325]

Vor dem Hintergrund dieser Theorie der Gefühle, schlägt Lahno vor auch Vertrauen als ein Gefühl aufzufassen. Als ein solches bestimmt auch Vertrauen wie ein Akteur seine Situation, und insbesondere seinen Interaktionspartner wahrnimmt. Insbesondere manifestiert sich Vertrauen als Emotion in einem Punkt, der im Zentrum der folgenden Diskussion stehen soll, und an dem wir das entscheidende Manko der Emotionstheorie festmachen werden: Die Vertrauende nimmt die andere Person unmittelbar als vertrauenswürdig wahr.[326]

Im Zusammenhang mit dieser Konzeption steht damit die Idee, dass sich Vertrauen nicht in

antizipieren und erwähnt zumindest, dass Vertrauen rationaler Kontrolle unterliege, wenn auch nicht „direkt“ (2001, 185). Dann aber bliebe es zumindest eine offene Frage, wie genau eine solches Moment rationaler Kontrolle in seine Emotionstheorie sinnvoll integriert werden könne.

323Die Begriffe *Gefühl* und *Emotion* benutze ich hier synonym.

324Ausdrücklich orientiert sich Lahno an der aristotelischen Definition von Emotionen: „Gefühle sind jene Dinge durch die über Veränderung Menschen dazu veranlasst werden sich in ihren Urteilen zu unterscheiden, und die von Schmerzen und Freude begleitet sind“ (rhet. 1378a).

325Lahno 2002, 205.

326Explizit in: ebd., 217, 219. Damit steht er auch mit Karen Jones‘ und Jessica Millers Emotionstheorien des Vertrauens im Einklang.

der Reflexion über mögliche Gründe für das Vertrauen zeigt. Vertrauen in eine Person zeige sich darin, dass diese *unmittelbar* als vertrauenswürdig empfunden werde, ohne dass der Vertrauende bewusst Gründe abwägt, die dafür oder dagegen sprechen die Vertrauensperson so zu sehen. Insofern Vertrauen als Emotion verstanden werden soll, gilt hier grundsätzlich: „[it does] not require the mediation of reason".[327] Damit scheint Lahno eine Konzeption von Vertrauen vorzulegen, die den spontanen und nicht-reflektierten Charakter, den Vertrauen haben kann, ins Zentrum rückt. Allerdings werden wir aufzeigen, dass die Emotionstheorie Lahnos an einer anderen Stelle ein entscheidendes Problem zu haben scheint.

Oben haben wir uns bereits ausgiebig mit Bemerkungen Lahnos zur Frage nach dem kognitiven Charakter von Vertrauen auseinandergesetzt. Wir haben versucht aufzuzeigen, dass seine Argumente gegen eine kognitivistische Position nicht wirklich überzeugen können. In Lahnos Buch *Der Begriff des Vertrauens* findet diese Diskussion noch bevor er seine Emotionstheorie einführt statt. Mit der Emotionstheorie aber ist Lahno eindeutig auf ein nicht-kognitivistisches Verständnis von Vertrauen festgelegt, und scheint damit vor grundlegende begriffliche Schwierigkeiten gestellt. Um dies genauer zu sehen, schauen wir uns die Diskussion eines konkreten Beispiels von Lahno an. Lahno beschreibt den Fall, in dem Julia unsterblich in Romeo verliebt ist, zugleich aber weiß, dass es sich bei Romeo um einen bekannten Heiratsschwindler handelt. In diesem Falle nun, so Lahnos Beurteilung, könne Julia durchaus Romeo vertrauen, während sie gleichzeitig der Überzeugung sei, dass Romeo nicht vertrauenswürdig ist.

Dass dies für Lahno keinen unmöglichen Widerspruch darstellt, hängt dabei mit seiner Emotionstheorie zusammen. Vertrauen stellt demnach ja allein eine Form des *unmittelbaren Wahrnehmens* einer anderen Person als vertrauenswürdig dar: „Julia urteilt nicht, dass Romeo vertrauenswürdig ist, sie *erlebt* es."[328] Der Lahnoschen Konzeption folgend ergibt sich deshalb, dass Vertrauen als Emotion unberührt *neben* reflektierteren Überzeugungen stehen kann, so wie im Falle Julias: Sie könne einerseits von der mangelnden Vertrauenswürdigkeit von Romeo überzeugt sein und ihm *trotzdem* vertrauen. Dieser Zusammenhang wird von Lahno auch mit Hilfe einer Analogie zu optischen Täuschungen erläutert. So nehme eine Person, die mit einer Illusion wie der Müller-Lyer-Täuschung konfrontiert sei, zwei Linien als unterschiedlich lang wahr, auch wenn sie weiß, dass dies nicht der Fall ist. Der Umstand, dass sie der begründeten Überzeugung ist, dass es sich um zwei gleich lange Linien handelt, verhindert nicht, dass sie dennoch die zwei Linien als unterschiedlich lang *sieht*. Die Situation sei parallel zu der, in der sich Julia befindet. Ihr Verstand

327Lahno 2001, 176.
328Lahno 2002, 219.

mag ihr zwar gewissermaßen sagen, dass Romeo nicht vertrauenswürdig sei, sie könne dennoch nicht anders als Romeo als vertrauenswürdig zu erfahren und ihm damit, gemäß Lahnos Theorie, zu vertrauen.

Das Problematische an dieser Position kommt deutlich zum Vorschein, wenn wir prägnant die Konsequenz dieses Ansatzes festhalten. Lahnos Idee folgend müssten wir Julia als eine Person beschreiben, die einerseits Romeo vertraut und ihn andererseits als nicht vertrauenswürdig ansieht. Dies aber, so hatten wir bereits zur Einleitung unseres Argumentes für die kognitivistische Position stark gemacht, scheint eine Moore'sche Aussage darzustellen.[329] Es scheint unmittelbar begrifflich widersprüchlich zu sein zu sagen, dass jemand einer Person vertraut, aber zugleich davon überzeugt ist, dass diese Person nicht vertrauenswürdig ist.[330]

Aus diesem Grund scheint mir auch die Parallele zu den optischen Täuschungen etwas schief gezogen zu sein. Für Wahrnehmungen mag es zwar stimmen, dass etwas auf eine bestimmte Weise zu sehen, noch nicht zwangsläufig bedeuten muss, dass man auch der *Überzeugung* ist, dass die Realität tatsächlich so beschaffen ist, wie sie erscheint. Das Beispiel in dem jemand weiß, dass er mit einer optischen Illusion konfrontiert ist, illustriert diesen Punkt. Vertrauen jedoch steht in einem oben dargelegten begrifflichen Zusammenhang zu einer Überzeugung oder Erwartung, dass der Vertrauenspartner vertrauenswürdig ist. Ohne dieses Moment einer Überzeugung kann man nicht wirklich plausibel von Vertrauen sprechen, und genau dieses Problem stellt sich für Lahno. Wenn er davon spricht, dass Julia nicht urteilt, dass Romeo vertrauenswürdig ist, sondern es vielmehr erlebt, scheint er ihr nicht wirklich Vertrauen zuzuschreiben, sondern vielmehr nur einen *Eindruck oder ein Gefühl* davon, dass jemand als vertrauenswürdig erscheint. Dass jemand für mich als vertrauenswürdig erscheint, bedeutet aber noch nicht, dass ich ihm tatsächlich vertraue. Stellen wir uns vor, ich betrachte ein Foto von Bernd Madoff und stelle fest, dass die Person auf dem Bild auf mich unmittelbar einen vertrauenserweckenden Eindruck macht. Ich erlebe Bernd Madoff in diesem Sinne also als vertrauenswürdig. Und doch: Sofern ich weiß, um wen es sich handelt, würde ich wohl kaum sagen, dass ich Vertrauen in ihn hätte.

Wir können diesen Punkt auch an einem weiteren Beispiel beleuchten. Erinnern wir uns dafür an den erwähnten von Judith Baker geschilderten Fall: Angenommen eine Freundin wird eines Verbrechens beschuldigt, versichert mir aber, dass sie unschuldig sei. Nun stellen wir uns vor, ich würde der Freundin Folgendes erwidern: „Du *wirkst* zwar für mich, als ob du unschuldig seist,

329Vgl. Kap. 7.2.

330Eine ähnliche Kritik scheint im Übrigen auch auf Karen Jones' Emotionstheorie des Vertrauens zuzutreffen. Siehe dazu: Cogley 2012, 33-36.

aber ich bin dennoch überzeugt, dass du schuldig bist“. Offensichtlich würde ich damit der Freundin nicht mein *Vertrauen* ihr gegenüber ausdrücken und vermutlich ihre Empörung auf mich ziehen. Meine Freundin erwartet nicht von mir, dass ich einen vertrauenswürdigen Eindruck von ihr habe. Sie erwartet, dass ich ihr glaube, sprich, ihr vertraue.

Das grundlegende Manko der Emotionstheorie, wie sie Lahno exemplifiziert, ist, dass sie den Unterschied nivelliert zwischen dem Vertrauen in eine Person und dem Eindruck, dass jemand vertrauenswürdig sei. Der Unterschied besteht in der kognitiven Überzeugung, die begrifflich mit Vertrauen einhergeht, jedoch nicht mit Emotionen, die sich nach Lahno in bloßen Eindrücken manifestieren.

Tatsächlich erkennt Lahno auch an, dass es problematisch ist, Vertrauen und Überzeugungen über Vertrauenswürdigkeit als *vollkommen* entkoppelt oder bloß zufällig übereinstimmend zu konzipieren. Er geht deshalb davon aus, dass in der Regel bestimmte kausale Zusammenhänge dafür sorgen, dass beide miteinander harmonieren. Dass Vertrauen und Erwartungen in der Regel zusammengehen, stelle damit keinen begrifflichen Zusammenhang dar, es sei „vielmehr eine Folge der Gesetzmäßigkeiten, denen unser bewusstes Leben unterliegt, und insofern im weitesten Sinne ein kausaler Zusammenhang.“[331] In Bezug auf das Beispiel mit Julia und Romeo merkt Lahno deshalb auch an, dass es sich um einen vermutlich seltenen Fall handele. Es bleibt aber damit dabei, dass Lahno davon ausgeht, dass Vertrauen und Überzeugung *begrifflich* unabhängig voneinander sind.[332]

Nur eine weiter verschärfende Folge der Lahnoschen Konzeption von Vertrauen als Emotion besteht darin, dass bestimmte Selbstverständlichkeiten in Bezug auf Vertrauen nicht erklärbar erscheinen. Insofern Lahno hier Vertrauen lediglich als eine emotionale Einstellung versteht, die man nun einmal hat oder nicht hat und die logisch distinkt neben widersprechenden epistemischen Einstellungen bestehen kann, stellt sich etwa die Frage, wie wir das Vertrauen anderer sinnvoll *kritisieren* können. An Lahnos Analogie zur optischen Täuschung wird dieser Punkt besonders deutlich: Einer Person, die eine optische Täuschung als eine solche *erlebt*, kann man dies nicht sinnvollerweise vorwerfen. Es scheint aber eine unproblematische

331Lahno 2002, 220.

332Der problematische Punkt von Lahno kommt lediglich noch schärfer zum Ausdruck, wenn man ein weiteres Theorem Lahnos hinzunimmt. So plädiert In Abgrenzung zur Rational-Choice Theorie, die Vertrauen wie gesehen als rationale Entscheidung auf der Basis kognitiver Überzeugungen konzipiert, argumentiert Lahno, dass die Verhältnisse exakt umgekehrt seien: Wir verweisen auf Vertrauen, um bestimmte Erwartungen zu erklären (ebd., 138-141). Diese Aussage wird zwar zuweilen von ihm auch relativiert – als nur in manchen Situationen, oder nur in bestimmten Grade gültig - wirft aber dennoch grundlegende Fragen danach auf, wie sich für Lahno das Verhältnis von Vertrauen und epistemischen Gründen darstellt.

Selbstverständlichkeit zu sein, dass wir eine Person für ihr Vertrauen tadeln können, wenn uns dieses etwa als naiv vorkommt. Oben haben wir darauf hingewiesen, dass eine solche Form von Kritik möglich ist, da sich die Vertrauende selber immer schon als unter dem Anspruch auf Situationsangemessenheit stehend versteht. Indem man sie tadelt wird sie an diesem selber gemessen. Lahnos ins A-Rationale abrutschende Konzeption hingegen hat Schwierigkeiten diesem Aspekt Rechnung zu tragen. Aus diesen Gründen stellen Emotionen im Lahnoschen Sinne keine geeignete Kategorie zur Konzeption nicht-reflexiver Vertrauensformen dar.

8.2 Die Tugendtheorie des Vertrauens

Einen anderen Strang in der Vertrauensliteratur, in dem auf die Reflexionslosigkeit von Vertrauen abgehoben wird, findet man in Miranda Frickers einflussreicher Monografie *Epistemic Injustice*.[333] Das zentrale Anliegen des Buches besteht in der Einführung des Konzeptes der epistemischen Ungerechtigkeit. Damit ist das Phänomen gemeint, dass bestimmten Sprechern, insbesondere solchen marginalisierter Minderheiten, aufgrund von Vorurteilen weniger Glaubwürdigkeit zugeschrieben wird, als ihnen gerechtfertigterweise zukommen sollte. Im Zuge dessen, und dieser Punkt wird im Folgenden im Vordergrund stehen, entwickelt Fricker eine tugendepistemologische Zeugenschafts- bzw. Vertrauenstheorie. Dabei stellt Fricker Analogien zwischen Vertrauen und Tugenden heraus, und zwar mit den zentralen Ziel, Vertrauen in seiner nicht-reflexiven Form verständlich zu machen. Wie wir gleich noch genauer aufzeigen werden, geht mit Tugenden auch eine Form des *unmittelbaren* Wahrnehmens statt eines ausdrücklichen Überlegens einher. Über die Emotionstheorie hinaus erlaubt es die Tugendtheorie des Vertrauens aber, dieses Wahrnehmen nicht als einen bloßen Eindruck, sondern als ein qualifiziertes Urteil des Vertrauenden zu verstehen.

Um ein Missverständnis gleich zu Beginn auszuräumen: Fricker will natürlich nicht sagen, dass Vertrauen stets eine Tugend *ist*. Vertrauen kann zwar zu einer Tugend ausgeformt werden, und ein Anliegen des Buches von Fricker ist es auch tugendhaftes Vertrauen zu propagieren. Darüber hinaus will sie Vertrauen aber nicht mit einer Tugend gleichsetzen, sondern ist daran interessiert strukturelle Parallelen zwischen beiden hervorzuheben. Das insbesondere neo-aristotelisch geprägt Tugendkonzept dient ihr insofern vielmehr als instruktives Modell, mit dem sich auch Vertrauen besser verstehen lassen könne.

333Fricker 2007.

Dabei ist ein entscheidender Punkt, den Fricker verstehen will, der, wie Vertrauen zu einem großen Teil unreflektiert und spontan ablaufen kann, ohne, dass dies hieße, dass es einfach „blind" vollzogen würde. Schließlich stelle es eine alltagsphänomenologische Tatsache dar, dass sich ein wesentlicher Bestandteil unseres Vertrauens in präreflexiver, automatisierter Form vollziehe. Entsprechend kritisiert Fricker (primär in Bezug auf Sprechervertrauen) die Vorstellung, Vertrauen basiere stets auf einem inferentiellen Urteilsprozess: „It simply does not match our everyday phenomenology of informal testimonial exchange, which presents learning something by being told as distinctly *un*-laborious and spontaneous."[334] Zugleich will sie aber zeigen, dass dies nicht heißt, dass wir damit einfach nur a-rational vertrauen. Auch in seiner präreflexiven Form könne man mehr oder weniger rational vertrauen und sei als Akteurin auch für die Rationalität des eigenen Vertrauens verantwortlich.

Um die grundlegende Idee Frickers zu erläutern, werden wir an dieser Stelle zunächst kurz darlegen, was sie unter einer Tugend versteht. Oben haben wir im Zusammenhang mit Vertrauenswürdigkeit bereits kurz erläutert, was eine Tugend im Wesentlichen ausmacht: Es handelt sich demnach um eine evaluativ positive Charakterdispositionen, die sich darin manifestiert, dass eine tugendhafte Person auf verlässliche Art aus guten Motiven handelt. Darüber hinaus stellt Fricker als entscheidenden Aspekt von Tugenden heraus, dass der Tugendhafte nicht erst ausdrücklich bei sich über die Abwägung von Gründen und Gegengründen darauf schließen muss, was richtigerweise zu tun ist. Vielmehr fasst er unmittelbar die relevanten Umstände der konkreten Situation, in der er sich befindet, als motivierende Handlungsgründe auf. Betrachten wir zur Illustration dieses Gedankens die Tugend der Hilfsbereitschaft: Wenn ein tugendhaft hilfsbereiter Mensch einer hilfsbedürftigen Person begegnet, so Fricker, dann überlegt er gar nicht erst, was zu tun ist, sondern sieht in der Hilfsbedürftigkeit des Anderen *unmittelbar* einen aktivierenden Grund zu helfen. Diese Idee geht insbesondere auf Überlegungen von John McDowell zurück, der festhielt: „The view of a situation which he [the virtuous person; J.W.]

334Fricker 2007, 63. Die Problemmotivation ist bei ihr allerdings zweifach etwas anders gelagert als bei uns. Was zum einen den Anspruch angeht, dass Vertrauen nicht einfach rein willkürlich sei, akzentuiert Fricker dies vorrangig als eine moralische Aufgabe: Sie will verstehen, was es heißt, möglichst vernünftig zu vertrauen, da sie vernünftiges Vertrauen als ein gerechtigkeitsförderndes Ziel ausgibt. Wir haben hingegen oben den Anspruch auf Situationsangemessenheit als eine begriffliches Merkmal verdeutlicht. Zum anderen versteht Fricker den Aspekt, dass man zuweilen nicht darüber nachdenken muss, ob man vertraut, als eine phänomenologische Tatsache unseres Alltagsvertrauens. So kritisiert sie in Bezug auf Sprechervertrauen die Vorstellung nach dem Modell expliziten inferenziellen Schließens folgendermaßen: „It *simply does not match our everyday phenomneology* of informal testimonial exchange, which presents learning something by being told as distinctly *un*-laborious and spontaneous" (ebd.: 63; Hervorhebung von mir). Auch hier haben wir vielmehr einen stärkeren Punkt im Auge: Dass Vertrauen zuweilen unreflektiert abläuft, stell nicht nur eine *akzidentelle* Beobachtung darüber dar, wie unser Vertrauen nun einmal faktisch ist. Vielmehr scheint der präreflexive Charakter ein (zumindest in manchen Kontexten) *wesentliches* Merkmal von Vertrauen zu sein.

arrives [...] is one in which some aspect of the situation is seen as constituting a reason for acting in some way; this reason is apprehended, not as outweighing or overriding any reasons for acting in other ways [...], but as silencing them.“[335]

Tugenden haben damit eine Art sensitive und eine motivationale Dimension, die sich aber nicht einfach voneinander trennen lassen. Wie gesagt, tugendhaft zu sein bedeutet, die Umgebung auf eine bestimmte Art und Weise *aufzufassen.* Es handelt sich um eine Art sensitiv gegenüber relevanten Handlungsgründen zu sein. Der Tugendhafte *erkennt* aber nicht bloß die relevanten Gründe in einer Situation, sondern er macht sie sich als *unmittelbar motivierende* Handlungsgründe zu eigen.[336] Qua seiner Tugendhaftigkeit ist die Art wie er seine Situation auffasst intrinsisch mit motivierenden Handlungsgründen verbunden. Diese Idee steht auch im Zentrum von Miranda Frickers Tugendverständnis. Eine Tugend gehe demnach mit einer Form von „moral perception“ einher, die zugleich unmittelbar Gründe gibt und intrinsisch motiviere.[337] Ein vermittelndes Element über einen *ausdrücklichen* praktischen Schluss ist für den Tugendhaften demnach nicht nötig. Eine Person, die erst explizit auf die richtige Handlung schließen muss, ist zwar ebenso wie der Tugendhafte praktisch am guten Handeln orientiert – beide zielen also darauf ab, das Gute zu tun. Allein der Tugendhafte jedoch hat die feste Disposition dazu erworben, *spontan* die Gründe für das gute Handeln zu erkennen und nach ihnen zu handeln. Insgesamt bringt der Tugendhafte also zwei Momente zusammen, die für unsere Diskussion von Interesse scheinen: Er handelt sowohl *spontan* als auch aus den *richtigen Gründen.*

Miranda Fricker listet darüber hinaus noch zwei weitere Punkte auf, die für Tugenden eine zentrale Rolle spielen. Dies ist zum einen der Umstand, dass das praktische Urteil des Tugendhaften nicht kodifizierbar ist. Es ließe sich also nicht als ein Regelfolgen des Tugendhaften beschreiben, da sich letztlich keine zutreffende generische Regel dafür angeben ließe.[338] Zum anderen führt Fricker aus, dass Tugenden immer auch eine emotionale Dimension haben. Diese beiden Punkte blenden wir hier jedoch aus, da es uns vordergründig darum geht zu verstehen, wie die Reflexionslosigkeit von Vertrauen verstanden werden kann, ohne, dass man sich das Manko des A-Rationalismus einkauft, wie es etwa für die Emotionstheorie zutrifft.

335McDowell 1979, 335.

336Nach einer bestimmten Lesart der aristotelischen Tugendtheorie, der auch Miranda Fricker folgt, besteht hier ein begrifflicher Zusammenhang zwischen der perzeptiven und der motivierenden Dimension von Tugenden. Diese Lesart macht etwa John McDowell (1979) stark: „the requirement imposed by the situation, and detected by the agent's sensitivity to such requirements, must exhaust his reason for acting as he does“. (ebd., 332). Der Tugendhafte nimmt in der Darstellung McDowells die Gründe gegen die tugendhafte Handlung gar nicht wahr. Aus Sicht des Tugendhaften sind die Gegengründe vielmehr unmittelbar *ausgeblendet* („silenced“) (ebd., 335).

337Fricker 2007, 76f.

338Fricker 2007, 74-76. Siehe zu diesem Punkt auch: Anscombe 2000, §33; McDowell 1979, 336-248.

Die obige Diskussion hat bereits vor Augen geführt, inwiefern eine Vorstellung von Vertrauen als einer Tugend die obige Schwäche der Emotionstheorie umgeht. So gehen Tugenden nicht bloß mit einem spontanen Eindruck der Umgebung einher, wie es für Emotionen im Sinne Lahnos der Fall ist, sondern vielmehr mit einem spontanen *Urteil*. Tugenden bieten, so Fricker „a model for non-inferential judgment.“[339] Dies ist der entscheidende Aspekt, in dem Tugenden eine aufschlussreiche Analogie zu reflexionslosen Vertrauensformen bereitzustellen versprechen.

Allerdings bleiben mit Frickers Vorstoß noch einige offene Fragen im Raum. Wir haben bereits deutlich gemacht, dass Frickers These nicht darin besteht zu sagen, dass Vertrauen *per se* eine Tugend ist. Dies liegt daran, dass man auch auf nicht-lobenswerte, sondern etwa naive Weise vertrauen kann, wohingegen Tugenden als lobenswerte Charaktereigenschaften definiert sind. Vielmehr diente das Konzept der Tugend Fricker als desiderables Leitbild: Vertrauen kann und sollte zu einer solchen Tugend ausgeformt werden. Damit stellen sich aber insgesamt noch zwei Fragen bezüglich ihres Ansatzes, an die wir im Folgenden anknüpfen werden. Erstens stellt sich die Frage nach einer allgemeinen Charakterisierung präreflexiven Vertrauens. Vertrauen kann zwar zu einer Tugend ausgebildet werden, muss es aber nicht zwangsläufig. Wie aber lassen sich dann allgemeingültig reflexionslose Vertrauensformen charakterisieren? Zweitens bleiben noch die oben aufgeworfenen Fragen offen: Wie genau sind die potenziellen Lernprozesse der Vertrauenden zu verstehen, sowie der Umstand, dass wir Vertrauende oder Nicht-Vertrauende für ihre Einstellung verantwortlich machen können?

Der Vorschlag, den wir im Folgenden entwickeln werden lautet, dass präreflexive Vertrauensformen in einem bestimmten Sinne als *habitualisierte Praktiken* aufzufassen sind. Dies verspricht eine plausible Kategorie bereitszustellen, mit der auch der Prozess des Erwerbs unreflektierter Vertrauensformen beschrieben werden kann. Dafür werden wir uns im folgenden Unterkapitel auch mit dem Prozess des Lernens und von Habitualisierungen auseinandersetzen.

8.3 Vertrauen und Misstrauen als habitualisierte Praxis

Die Idee, die wir in diesem Kapitel erläutern werden lautet also, dass sich Vertrauen und Misstrauen zu einem bedeutenden Teil in Form von *habitualisierten Praktiken* darstellen. Die verschiedenen Praktiken des Vertrauens haben im Alltagsleben einen essentiellen und ubiquitären

339Fricker 2007, 72.

Stellenwert. Es gibt wohl kaum eine soziale Tätigkeit, die sich nicht auch als eine Praxis des Vertrauens beschreiben ließe. Denken wir etwa daran, wenn wir etwa Geld von jemanden annehmen, im Vertrauen, dass dieses nicht gefälscht ist; wenn wir die Zeitung lesen, im prinzipiellen Vertrauen, dass gewissenhaft über Ereignisse in der Welt berichtet wird; oder überhaupt, wenn wir uns angstfrei mit unseren Mitmenschen im öffentlichen Raum bewegen etc. All diese eingewöhnten Praktiken laufen, solange sie nicht durch konkrete Anlässe irritiert werden und reibungslos funktionieren, größtenteils unreflektiert ab. Das in ihnen inkorporierte Vertrauen zeigt sich in unseren quasi-natürlichen Handlungsweisen und bleibt in der Regel unthematisch.

Im Folgenden werden wir ausführlicher ein bestimmtes Praxiskonzept entwickeln, das verspricht die Form präreflexiven Vertrauens akkurat zu erfassen. Im Allgemeinen spielt der Begriff der Praxis in der philosophischen und sozialwissenschaftlichen Diskussion in verschiedenen Traditionslinien eine zentrale Rolle und hat daher auch recht unterschiedliche Bedeutungen und Stellenwerte zugeschrieben bekommen.[340] Die folgende Fassung des Praxisbegriffes anhand einiger Merkmale kann die Breite der verschiedenen Praxisverständnisse nicht erfassen und repräsentiert vielmehr ein spezielles Praxisverständnis, das passgenau auf unser Anliegen gemünzt ist. Dieses Anliegen besteht darin, Praktiken als eingewöhnte und habitualisierte Tätigkeitsformen auszuweisen, die stillschweigend und überlegungsfrei vollzogen werden.[341]

Man beachte, dass „Praxis“ dabei nicht ausschließlich im Sinne einer sozial geteilten Praxis zu verstehen ist.[342] Der hier veranschlagte Praxisbegriff kann sich vielmehr auch auf individuell eingewöhnte Praktiken beziehen. Auf Individuen bezogen könnte man gleichbedeutend auch von

340Zur Konjunktur des Praxis-Konzeptes, siehe beispielsweise: Alkemeyer et.al. (2016); Schäfer (2016); Cetina/Schatzki (2001).

341Wir orientieren wir uns bei der Entwicklung dieses Verständnisses maßgeblich an vereinzelten Wittgensteinschen und pragmatistischen Überlegungen. Insgesamt hat das entwickelte Praxisverständnis mit Sicherheit eine große Nähe zum pragmatistischen Verständnis zum Konzept der Gewohnheit (*habit*). Siehe etwa die Charakterisierung John Deweys einer solchen: „The word habit may seem twisted somewhat from its customary use when employed as we have been using it. But we need a word to express that kind of human activity which is influenced by prior activity and in that sense acquired ; which contains within itself a certain ordering or systematization of minor elements of action ; which is projective, dynamic in quality, ready for overt manifestation; and which is operative in some subdued subordinate form even when not obviously dominating activity.“ (Dewey 2007, 31).

342Die Idee sozial geteilter Praktiken ist sicherlich wichtig, etwa auch um aus soziologischer Sicht die Kohäsion größerer sozialer Kollektive zu erklären. Unter Bezugnahme darauf ließe sich zudem etwa auch ein Konzept wie das des Vertrauensklimas erläutern (Vgl. Hartmann 2011, 296-300).Unter einem guten Vertrauensklima innerhalb einer sozialen Gruppe, ließe sich etwa ein Zustand verstehen, in dem sozial geteilte Vertrauenspraktiken mit einem hohen Generalisierungsgrad bestehen. Die Rede vom „Generalisierungsgrad“ ist dabei so zu verstehen, dass wir es nicht nur mit isolierten Vertrauensbereichen zu tun haben, dass also nicht nur in spezifischen Hinsichten spezifischen Instanzen vertraut wird, während anderen etwa misstraut wird (Vgl. das Konzept der *zones of trust* in: Walker 2006, 84f.). Vielmehr bedeutet ein hohes Vertrauensklima, dass sich über weite Bereiche erstreckende, stabile Vertrauenspraktiken ausgebildet haben. Bei einem hohen Vertrauensklima vertrauen sich Personen unreflektiert in Bezug auf mannigfaltige Sachverhalte. Diese verschiedenen Vertrauenspraktiken stehen dabei natürlich auch in einem konsistenten, holistischen Geflecht zueinander. Im Normalfall sind sie miteinander kohärent oder begründen sich gar gegenseitig.

einer persönlichen Disposition reden. Unten, wenn wir uns erneut spezifisch dem interpersonellen Vertrauen zuwenden, werden wir auch von einem individuellen *Habitus* sprechen, wenn es darum geht zu betonen, dass man sich mit seinen habitualisierten Einstellungen auch zu anderen Personen verhält. An dieser Stelle aber bleiben wir allgemeiner beim Begriff der Praxis.

Ein besonderes theoretisches Anliegen besteht zudem darin, intelligibel zu machen, wie Vertrauensformen als unüberlegt aber nicht einfach nur als blind vollzogen gedacht werden können. Und damit zusammen hängt ebenso die Beantwortung der Frage, wie es zu denken ist, dass Vertrauende oder Misstrauende für ihr Vertrauen oder Misstrauen verantwortlich sind.

8.3.1 Praxis als kognitiv strukturierte Tätigkeitsform

Grundlegend gesprochen handelt es sich bei einer Praxis nicht um eine konturlose Menge an Einzelakten, sondern vielmehr um etwas *Allgemeines*, was sich *in verschiedenen Einzelakten manifestiert*. Zur Praxis gehört also, dass sie sich in konkreten Akten ausdrückt. Die Praxis selber aber ist das übergreifende Allgemeine, dass sich in diesen wiederkehrenden Einzelakten zeigt.[343] Weniger definitorisch verstanden, sondern vielmehr genealogisch gewendet, besteht zudem die *Aneignung* einer solchen Praxis in der wiederholten und störungsfreien Ausübung von Einzelakten, die diese Praxis instantiieren.

Es ist noch einmal wichtig zu betonen, dass Praktiken in diesem Sinne nicht notwendigerweise als *sozial geteilte* Praktiken zu verstehen sind. Es handelt sich also nicht zwangsläufig um Praktiken, die von mehreren Mitgliedern eines sozialen Kollektivs regelmäßig ausgeführt werden. Der Begriff kann vielmehr auch im Sinne einer individuellen Praxis verstanden werden, einer Praxis also, die nur eine Person gewohnheitsmäßig betreibt. Das soll nicht bedeuten, dass auch individuelle Praktiken soziale Voraussetzungen haben und in andere soziale Praktiken eingebettet sind. Diese Zusammenhänge zwischen der sozialen Dimension und der individuellen

343 Praktiken sind damit auch nicht einfach ein Bündel oder die Generalisierung von Einzelakten, da solche Einzelakte immer schon als Instanziierung einer allgemeinen Praxis verstanden werden. Siehe etwa die Erläuterung von Michael Thompson diesbezüglich: „In describing a practice that prevails among some class of agents we say "what they do", in some sense, and not what any or many or most or all of them "are doing," "have done" or "will do." The thing characterized thus falls outside the order of individual events into which individual actions are fitted" (Thompson 2008, 156). Damit erweist sich das Verhältnis zwischen Allgemeinen und Konkreten bei Praktiken als verwickelt: Einerseits bestehen Praktiken zwar nur, insofern sie manifestiert werden. Diese Manifestationen wiederum werden immer schon im Lichte eines allgemeinen Praxisverständnisse *als* Manifestationen einer bestimmten Praxis verstanden. Deswegen hält Mathias Haase fest: „Eine Praxis ist eine ‚Form' bestimmter Art, sie existiert nur als Form einer Multiplizität von Exemplaren" (Haase 2007, 240).

Dimension von Praktiken werden wir an dieser Stelle jedoch nicht beleuchten.

Im Zentrum einer Praxis stehen spezifische „Regeln“ oder „Normen“ durch die diese konstituiert ist. Diese geben vor, was in einer spezifischen Art von Situation gemäß der Praxis zu tun ist. Hierbei sind unter dem Begriff „Regel“ oder „Norm“ keinesfalls explizit kodifizierte Direktiven gemeint. Gemäß eines weiten und pragmatistisch verwurzelten Verständnisses des Regel- bzw. Normenbegriffes, beschreiben diese lediglich die eine Praxis auszeichnenden Verfahrensweisen, die bestimmen, was in einer Situation zu tun richtig oder angemessen ist und an denen sich die Teilnehmenden der Praxis implizit normativ orientieren. Wenn eine Person etwa in der Praxis des Autofahrens eingeübt ist, dann weiß sie, was in einer gegebenen Situation zu tun das richtige und angemessene ist. Wenn sie etwa auf eine Kreuzung zufährt, ist sie intuitiv damit vertraut, dass sie den Verkehrsteilnehmern von rechts in der Regel Vorfahrt zu gewähren hat.

Dass sich die Teilnehmenden an den praxiskonstitutiven Normen orientieren, soll hierbei mehr heißen, als dass sich eine *bloße* Regelmäßigkeit im Verhalten der Teilnehmenden zeigt. Vielmehr *folgen* die Teilnehmenden einer Praxis deren impliziten Regeln, das heißt: Sie orientieren sich *intentional* in ihrem Handeln an diesen Regeln. Den Regeln zu folgen bedeutet aus der Sicht einer Akteurin, dass sie als handlungsleitender Maßstab anerkannt sind, dass sie also als situativ angemessen betrachtet werden.[344] Der Pragmatist James Wallace hebt diesen Punkt unter Bezugnahme auf Frederick Will hervor, indem er „Normen“ im hier einschlägigen Sinne definiert als „learned activity that is taken to be an appropriate way to proceed in a certain domain.“[345] Die Autofahrerin etwa, die an der Kreuzung den Verkehrsteilnehmern Vorfahrt gewährt, versteht ihr Handeln als der Situation angemessen. Dies heißt wiederum gerade nicht, dass den Handelnden zwangsläufig thematisch sein muss, was sie tun und welcher Regel sie folgen. Eine in einer solchen Praxis verkörperte Regel kann zwar prinzipiell expliziert und in kodifizierter Form wiedergegeben werden. Dann gibt man jedoch nur deren *kodifizierte Seite* wieder, die Will als die „manifest form“ einer Norm bezeichnet. Ihr Charakter als stillschweigend mitlaufender Maßstab zur Handlungsorientierung wird damit nicht eingefangen.

Dabei setzt einer Regel zu folgen wesentlich eine Form von *Vertrautheit* mit dieser voraus. Man kann einer Regel trivialerweise nur folgen, sofern man diese kennt. Wie gesagt ist dieses Wissen den Teilnehmenden aber nicht zwangsläufig in einem ausdrücklichen Sinne bewusst,

344So hält Wallace fest: „Knowledge of how to do something is normative; it is knowledge of how to do it properly, knowledge of better rather than worse ways of doing things“ (Wallace 2009, 11). Die Regel stellt damit auch einen praxisinternen normativen Maßstab für die Handlungen dar. Wer beim Schachspielen den Läufer senkrecht über das Feld zieht, der macht etwas falsch. Gemessen an der üblichen Form des Schachspielens ist sein Handeln defekt.

345Ebd., 33.

sondern zeigt sich vielmehr als in ihren eingespielten Handlungsweisen verkörpert. James Wallace führt deshalb den Begriff des *praktischen* Wissens ein. Damit ist ein Wissen gemeint, das sich wesentlich im habitualisierten praktischen Umgang handelnder Akteurinnen zeigt, die sich implizit an bestimmten Regeln orientieren. Folglich definiert Wallace Praktiken als „activities guided by a [...] body of practical knowledge."[346] Vor dem Hintergrund dieser Erläuterungen wird im Übrigen auch besonders deutlich, inwiefern der Begriff der Praxis hier nicht als strenger Gegenbegriff zur Theorie gedacht werden kann. Vielmehr sind auch Praktiken im eben erläuterten Sinne von einer kognitiven Struktur, einem praktischen *Wissen* getragen.

Auch Vertrauen in Form solcher habitualisierten Praktiken lässt sich nun als durch Regeln strukturiert verstehen. Spezifische Praktiken des Vertrauens oder Misstrauens folgen spezifischen impliziten „Regeln", die angeben, in welchem Kontext welchen Personen zu vertrauen oder zu misstrauen ist. Eine ausgewachsene Praxis des Vertrauens, in eine bestimmte Expertengruppe X beispielsweise, lässt sich in diesem Sinne beschreiben als durch die verhaltensleitende Regel konstituiert, dass Expertinnen des Typus X in der Regel vertraut wird. Diesem Verständnis nach können etwa auch die in der Diskussion von Fricker fokussierten *Vorurteile* als praxisleitende Regeln aufgefasst werden. Eine Akteurin mit Vorurteilen gegenüber bestimmten Personengruppen (Fricker verwendet auch den Terminus des *sozialen Typus*[347]) lässt sich von diesen auch, zumeist unreflektiert, in ihren spontanen Neigungen zu vertrauen oder zu misstrauen leiten. In diesem Sinne fungieren Vorurteile wie die eben beschriebenen praxisinternen Regeln. Anhand dieser Erläuterung wird auch noch einmal deutlich, dass Regeln hier nicht notwendigerweise als moralisch ausgezeichnet zu verstehen sind, sondern diese im Gegenteil auch eine unmoralische Qualität haben können.

Weiterhin sei noch angemerkt, dass Praktiken, zumindest dem Anspruch nach, je gewissen Zwecken dienen. Eine Praxis würde nicht ausgeführt werden, würde sie als zwecklos erachtet. Vielmehr werden Praktiken stets im Lichte eines oder mehrerer für sie konstitutiver Zwecke verfolgt. Genealogisch gewendet lassen sich Praktiken auch als Lösungen auf ein bestimmtes Problem sehen. Dies definiert ihren Zweck. Eine Praxis des Misstrauens gegenüber bestimmten Personengruppen beispielsweise, ließe sich verstehen als die Antwort auf die wiederholte problematische Erfahrung, dass das Vertrauen in diese enttäuscht wurde. In dieser Erfahrung wird deutlich, dass die vorhergehende Praxis des Vertrauens nicht mehr ihre eigentlich eigenen Zwecke

346 Wallace 2009, 11. Im Original heißt es: „activities guided by a shared body of practical knowledge". Dies bezieht sich speziell auf *soziale* Praktiken, deswegen wird das Wissen als *geteilt* charakterisiert.
347 Fricker 2007, 71.

(etwa die Vorzüge kooperativer Beziehungen) erfüllen kann. Vor diesem Hintergrund kann es sich geradezu aufdrängen eine Praxis des Misstrauens herauszubilden.

Wiederum gilt auch hier, dass den Teilnehmenden der Praxis die Natur ihres Zweckes nicht vollständig klar, und schon gar nicht ausdrücklich präsent sein muss. In vielen Kontexten wird vielmehr latent angenommen, dass Praktiken gewissen Zwecken dienen. Diese Annahme kann sich natürlich auch als trügerisch erweisen und durch reale Erfahrungen widerlegt werden. In diesem Fall zeigen sich akute Störungen oder Krisen der Praxis, die, wie eben angedeutet, zu einer Abänderung derselben drängen. Wenn eine Praxis nicht (mehr) ihren originären Zwecken dient, kann sie zudem auch ausdrücklich kritisiert werden.[348] In jedem Fall können sich die Teilnehmenden einer Praxis auch immer reflexiv zu dieser verhalten, diese gemessen an ihren eigentlichen Zwecken bewerten und gegebenenfalls abändern.

8.3.2 Der Automatizitätcharakter von Praktiken

Die praxisleitenden Regeln können expliziert werden, im Ausüben einer eingewöhnten Praxis wird diesen jedoch spontan und unreflektiert gefolgt. Diesen geht in der Regel kein explizites Nachdenken voraus, vielmehr werden sie gewissermaßen „unbedacht“ vollzogen. Diese Reflexionslosigkeit bezieht sich aber nicht nur auf die Vorstellung dessen, was man tut, sondern auch auf die möglichen Gründe, warum man dies tut: Insofern Akteurinnen eine habitualisierte Tätigkeit ausüben, sind sie davon entlastet ausdrückliche Überlegungen über die Gründe für ihre Handlungen zu reflektieren. Dieser Punkt macht etwa bei David Hume gerade den Kern seines Konzepts einer Gewohnheit aus, die eng mit unserem Praxiskonzept verwandt ist: „For wherever the repetition of any particular act or operation produces a propensity to renew the same act or operation, without being impelled by any reasoning or process of the understanding; we always say, that this propensity is the effect of *Custom*.“[349]

Zuweilen spricht man in der philosophischen und psychologischen Literatur zum Thema habitualisierten Verhaltens auch von dessen Automatizitätscharakter. Automatizität stellt dabei einen Cluster-Begriff dar, der ein Bündel an bestimmten Eigenschaften zusammenfasst: Etwa die Abwesenheit von bewussten Absichten, von bewusster Aufmerksamkeit bei der eigenen Tätigkeit

348Siehe hierzu ausführlicher: Jaeggi 2014.
349Hume 2007b, 32.

und eine gewisse Anstrengungslosigkeit.[350] Damit einher geht auch das Absehen von Reflexion über die Gründe des eigenen Handelns.

8.3.3 Lernprozesse und Verantwortlichkeit: Die Idee kontrafaktischer Sensitivität

Wie am Ende des vorletzten Abschnittes bereits angemerkt wurde, bedeutet der Umstand, dass Praktiken präreflexiv vollzogen werden, nicht, dass diese keiner kritischen Reflexion bzw. Abänderung unterzogen werden können. Praktiken können sich vielmehr im Zuge von Lernerfahrungen dynamisch entwickeln, insbesondere sofern sich praktische Probleme bei ihrer Ausübung stellen und sie nicht mehr reibungslos funktionieren. Dieser Aspekt ist zudem der Schlüssel, um nachzuvollziehen, inwiefern Vertrauenspraktiken zwar einerseits nicht bewusst vollzogen werden, sie dennoch nicht einfach einem a-rationalen, blindem Reagieren entsprechen.

Wir wollen uns nun genauer dem Verhältnis der in praktischen Zusammenhängen stehenden Akteurinnen und ihrer Verantwortlichkeit zuwenden. Die zentrale Frage wird sein, wie habitualisierte Verhaltensweise zwar einerseits durch ihren Automatizitätscharakter und ihre Reflexionslosgikeit gekennzeichnet sind, wir zugleich jedoch Akteurinnen für eingeschliffene Handlungsmuster verantwortlich machen können. Insbesondere werden wir dies als Gelegenheit nutzen, um zu zeigen, wie die häufig in der Vertrauensliteratur geäußerte Idee einer kontrafaktsichen Sensitivität hier hineinspielen kann.

Zunächst aber wollen wir das zugrundeliegende Problem mithilfe einer Textpassage aus den *Philosophischen Untersuchungen* von Ludwig Wittgenstein schärfer artikulieren. Anhand dieser lässt sich rekonstruieren, wie sich beim Versuch habitualisiertes Handeln zu verstehen ein theoretisches Dilemma präsentiert. Wittgenstein erörtert dabei den Bezug zwischen spontanen, habitualisierten Handlungsweisen und deren rechtfertigenden Gründe. Dabei ist er im Besonderen darum bemüht, habitualisiertes Handeln als ein reflexionsloses aber dennoch nicht einfach a-rationales Phänomen auszuweisen.[351] Wittgenstein schreibt:

350 Vgl. Douskos 2017, 1135.

351 Reflexionen zu habitualisierten Praxisformen finden sich bei Wittgenstein insbesondere in den hinteren Paragraphen der *Philosophischen Untersuchungen*. Darüber hinaus spielt dies auch in Wittgensteins Reflexionen aus *Über Gewissheit* (Wittgenstein 2015*)* eine zentrale Rolle. In den vergangenen Jahren ist insbesondere in diesem Zusammenhang auch der Einfluss des Pragmatismus auf Wittgensteins Spätphilosophie hervorgehoben worden (Vgl. dazu Philström 2012).

„- »Die Gewißheit, daß das Feuer mich brennen wird, gründet sich auf Induktion.« Heißt dies, daß ich bei mir schließe »Ich habe mich immer an einer Flamme verbrannt, also wird es auch jetzt geschehen«? Oder ist die frühere Erfahrung die *Ursache* meiner Gewißheit, nicht ihr Grund? Ist die frühere Erfahrung die Ursache der Gewißheit - das kommt auf das System von Hypothesen, Naturgesetzen an, in welchem wir das Phänomen der Gewißheit betrachten. Ist die Zuversicht gerechtfertigt?“ [352]

In dem Abschnitt geht es Wittgenstein um den Zusammenhang zwischen vergangenen Erfahrungen („Ich habe mich immer an einer Flamme verbrannt“) und der gegenwärtigen Gewissheit einer Person angesichts einer mit diesen Erfahrungen vergleichbaren Situation. Der Begriff der Gewissheit ist hierbei natürlich ein zentraler Begriff in Wittgensteins Spätphilosophie, der insbesondere in seiner Auseinandersetzung mit dem Skeptizismus in *Über Gewissheit* eine Rolle spielt.[353] Wir können hier keine ausführliche Diskussion des Begriffes und seiner vielfältigen Bedeutungsinhalte im Kontext von Wittgensteins Spätphilosophie vornehmen. Zentral ist zunächst, dass er solche Gewissheiten nicht als *bloß* kognitive Annahme versteht, sondern als Annahmen, die sich wesentlich in unseren spontan zeigenden praktischen Dispositionen manifestieren. An anderen Stellen bringt Wittgenstein auch deutlich in Bezug auf das Flammen-Beispiel zum Ausdruck, dass er hier die Gewissheit als eine Handlungsdisposition auffasst: „Nichts könnte mich dazu bewegen, meine Hand in die Flamme zu stecken, – obwohl ich mich doch *nur in der Vergangenheit* verbrannt habe.“[354] Dieser wesentlich praktisch leitende Charakter von Gewissheiten lässt diese also als Form dessen erscheinen, was wir oben mit James Wallace als praktisches Wissen bezeichnet haben.

Das Problem, dass wir nun anhand des Zitates motivieren werden, lautet: Inwiefern ist der unwillkürliche Charakter des Handelns aus Gewissheit mit der Verantwortlichkeit der handelnden Akteurin vereinbar? Schauen wir uns aber den Gedankengang des Zitates schrittweise an. Die Frage, die Wittgenstein stellt, ob ich „bei mir schließe“ dass die Flamme mich brennen wird, ist offenbar rhetorisch zu verstehen: Wir schließen normalerweise in solchen Fällen nicht ausdrücklich „bei uns“ aus vergangenen Erfahrungen auf aktuale Einstellungen oder Handlungen. Hier hebt

352 Wittgenstein 2003a, §325.

353 Vgl. Wittgenstein 2015. Gewissheiten stellen dabei Annahmen dar, die sich insbesondere durch ihre fundamentale Rolle in einem System von Überzeugungen auszeichnen. Im Kontext dieser Arbeit interessieren uns Gewissheiten aber nicht in ihrem epistemisch-kontextualen Status, sondern in ihrer vielmehr phänomenologischen Eigenart als Überzeugungen, über die Akteure keine bewusste Reflexionen anstellen, sondern, die sich vielmehr in ihrem spontanen Verhalten manifestieren – ohne, dass diese damit als Ausdruck eines blind reflexhaften, a-rationalen Verhaltens zu verstehen sind.

354 Wittgenstein 2003a, §472. An anderer Stelle betont er auch dessen notwendige affektive Seite, auf die wir hier nicht genauer eingehen können (ebd., §473). Insgesamt scheint Wittgenstein auch darum bemüht deutlich zu machen, dass dies für solche Phänomene wesentlich ist, und das man sie verfehlt, wenn man sie als theoretisches Schließen, oder auch nur als defizitäre Annäherung an ein solches Schließen versteht (vgl. ebd., §481).

Wittgenstein also ebenso auf den Automatizitätscharakter eingewöhnter Gewissheiten ab.

Trotz deren spontanen, unreflektierten Charakters, können diese nicht einfach nur als a-rationale Reflexe verstanden werden. Dieser Punkt scheint sich auch auf subtile Weise in dem Wittgensteinschen Paragraphen anzudeuten. So fragt Wittgenstein zunächst: „Oder ist die frühere Erfahrung die Ursache der Gewißheit – das kommt auf das System von Hypothesen, Naturgesetzen an, in welchen wir das Phänomen Gewißheit betrachten." Dass die Flamme mich nicht brennen wird, weiß ich entweder aus einem ausdrücklichen Schluss, der sich auf vergangene Erfahrungen bezieht. Dies scheint jedoch eine phänomenologisch abwegige Beschreibung zu sein, *wie* wir in der Regel in solchen Fällen urteilen. Dann aber legt sich die alternative Sichtweise nahe, dass mein Urteil als Ergebnis des Wirkens *bloßer*[355] Naturkausalität zu betrachten ist. Wittgenstein gesteht auch zu, dass dies eine mögliche Perspektive auf das Verhaltens der Person sein kann – eine Betrachtungsweise, die bestimmte Annahmen voraussetzt, nämlich einem „System von Hypothesen, Naturgesetzen". Als sozialempirisch interessierter Beobachter könnten wir beispielsweise aus den vergangenen Erfahrungen einer Person auf deren zukünftiges Verhalten als Wirkung schließen. Wittgenstein scheint an dieser Stelle also eine mögliche naturkausal geprägte Betrachtungsweise auf das Phänomen der Gewissheit zuzulassen.

Das bedeutet jedoch nicht, dass man damit den Anspruch erheben kann, das betrachtete Phänomen aus dieser Perspektive *vollständig* erfasst zu haben.[356] Vielmehr lässt Wittgenstein im weiteren Verlauf des Passus eine alternative Betrachtungsweise anklingen. Er stellt interessanterweise im direkten Anschluss die Frage: „Ist unsere Gewissheit *gerechtfertigt*?"[357] Der interessante Aspekt hieran, den ich hervorheben möchte, ist zunächst weniger die mögliche Antwortmöglichkeit auf diese Frage, sondern vielmehr die Tatsache, dass sie überhaupt an dieser Stelle von Wittgenstein aufgeworfen wird. Sie folgt nämlich auf den soeben betrachteten Hinweis, dass sich die Gewissheit auch naturkausal betrachten ließe. Die *Anwendbarkeit* der Frage nach einer *Rechtfertigung* deutet demgegenüber jedoch ein alternatives Verständnis an. Insofern wir uns nach einer Rechtfertigung für habitualisiertes Verhalten fragen, betrachten wir dieses offensichtlich

355Ich betone, dass es sich um *bloße* Naturkausalität handelt, da auch Gründe und Handlungen in einem *gewissen* kausalen Zusammenhang stehen. In der Handlungstheorie findet sich entsprechend die Idee, Kausalität als einen Genus zu verstehen, unter den verschiedene Spezies fallen. Eine Spezies der Kausalität, so die Idee, bezieht sich auf allein mechanische Prozesse. Zur Beschreibung des kausalen Zusammenhanges zwischen Gründen und Handlungen ist diese aber nicht geeignet (Vgl. Kietzmann 2019, 32-35).

356Dies zeigt sich auch in den nachfolgend von Wittgenstein verfolgten Reflexionen über den Unterschied zwischen der Ursache und dem Gegenstand von Furcht, die deutlich macht, dass in der bloß kausalen Beschreibung der hermeneutische Zugang zur Beschreibung der Perspektive des Handelnden nicht erfasst wird: „So ist das Gesicht, das uns Furcht, oder Entzücken, einflößt (der Gegenstand der Furcht, des Entzückens) darum nicht ihre Ursache, sondern – man könnte sagen – ihre Richtung" (Wittgenstein 2003a, §476).

357Ebd., §325.

nicht *bloß* naturkausal aus vergangenen Erfahrungen resultierend. Das Entscheidende, was die Frage nach der Rechtfertigung unterstellt ist, dass man auch für habitualisierte Einstellungen und Handlungen *verantwortlich* ist. Und man ist für diese verantwortlich, obwohl diese nicht auf ausdrücklichen Überlegungen des Akteurs beruhen.

Dass ein handelnder Akteur verantwortlich war, setzt also nicht voraus, dass er bewusst aus Gründen gehandelt hat. Es setzt aber auch nicht voraus, wie man meinen könnte, dass der Akteur dazu in der Lage wäre seine Gründe retrospektiv zu benennen. Zwar kann es sein, dass er, wie Wittgenstein es an einer anderen Stelle ausdrückt, sich auf seine Gründe „besinnt“[358]. Dies ist aber kein notwendiges Kriterium dafür, dass der Akteur für sein Verhalten verantwortlich ist. Dies ist nicht notwendig, da der Akteur, auch wenn ihm keine Gründe einfallen, trotzdem verantwortlicher Akteur war. Allein, insofern die *Frage* nach der Rechtfertigung seiner Handlung, die „Warum-Frage“, wie G.E.M. Anscombe es ausdrückt, *angemessen* ist, stellt er einen verantwortlichen Akteur dar.[359] Nur wenn die Warum-Frage, oder Wittgensteins Frage nach der Rechtfertigung, als *kategorial unangemessen* zurückgewiesen werden müsste, hätten wir es nicht mit einem verantwortlichen Akteur zu tun. Dies wäre etwa der Fall, wenn jemand seine Hand aufgrund eines Krampfes vom Feuer zurückzieht. Dies stellt einen bloß kausalen Vorgang vor, für den die agierende Person nicht verantwortlich ist und an die sich deshalb auch im Normalfall nicht sinnvoll die Frage richten ließe, ob ihre Handbewegung gerechtfertigt war.

In der Tat ist es selbstverständlicher Teil unserer moralischen Praxis, dass wir Akteurinnen nicht nur für Handlungen verantwortlich machen, die aus ausdrücklichen Überlegungen und als Ergebnis eines Entscheidungsprozesses vollzogen werden. Vielmehr machen wir Akteure mitunter auch für ihre habitualisierten Verhaltensweisen verantwortlich. Greifen wir etwa unseren Fall von oben auf: Zuweilen machen wir andere Personen für ihre durchaus habitualisierten und implizit wirksamen Vorurteile verantwortlich. Dies ist der entscheidende Impuls, den wir aus dem Wittgenstein‘schen Zitat herausziehen und vertiefen werden.

Dafür spitzen wir zunächst noch einmal genauer das theoretische Problem zu, das sich an dieser Stelle aufdrängt. Dieses stellt sich folgendermaßen dar: Wenn wir andere Personen verantwortlich machen, unterstellen wir, dass diese sich *hätten anders verhalten können*.[360] Diese Unterstellung, dass die andere Person hätte anders handeln können, ist nun intelligiblerweise

358 Wittgenstein 2003a, §475.

359 Anscombe 2000, §5.

360 Ich schließe hier allerdings Fälle von „politischer Verantwortlichkeit“ und ähnliche aus, für die ich diesen Punkt offen lasse. Hier geht es nicht um die Verantwortung, die sich allein aus dem Status oder der institutionellen Position einer Person heraus begründen.

angemessen, wenn wir uns die Handlung der anderen Person als Ergebnis einer *bewussten*, abwägenden Entscheidung vorstellen. Die Vorstellung, dass die andere Person sich entscheidet, impliziert immerhin, dass sie sich selber als eine Akteurin versteht, die zwischen verschiedenen, alternativen Handlungsoptionen zu wählen hat. Sie versteht sich also als eine Akteurin, die sich offenkundig angesichts verschiedener Handlungs*möglichkeiten* verhalten muss. Rätselhafter ist deshalb die Verantwortungszuschreibung im Falle habitualisierten Handelns, da sich dieses ja gerade dadurch auszeichnet, dass die Akteurin *ohne* Überlegung und bewusste Entscheidung handelt. In ihrer Perspektive präsentieren sich also gerade gar keine alternativen Handlungsmöglichkeiten, zu denen sie sich bewusst verhält. Inwiefern gleicht ihr Handeln dann aber nicht einem bloßen Naturgeschehen, für das sie nicht zu verantworten ist? [361]

Ein erster Antwortversuch auf diese Frage könnte folgendermaßen aussehen. Er besteht darin, zu sagen, dass Habitualisierungen insofern keine bloßen a-rationalen, bloß kausalen Prozesse darstellen, als es für die Akteurin stets *möglich* ist Abstand zu diesen zu gewinnen, diese explizit zu hinterfragen und daraufhin abzuändern. Die ohne Überlegung handelnde Akteurin ist demnach deshalb zu verantworten, da sie prinzipiell *hätte überlegen können*. Derart formuliert würde der Vorschlag uns jedoch ein neues Rätsel vorlegen: Was hieße es schließlich aus der Perspektive der Handelnden sich angesichts dieser rein hypothetischen Möglichkeit zu verhalten? Hier zeigt sich ein Problem, denn insofern sie habitualisiert handelt, ist sie sich dieser Möglichkeit ja gerade nicht bewusst. Sie handelt schließlich einfach, ohne Überlegung. Die hypothetische Möglichkeit anders zu handeln zu können, kommt *im Horizont der habitualisiert Handelnden* ja gerade nicht vor, sofern sie habitualisiert handelt. Würde sie hingegen ausdrücklich vorkommen, dann handelte die Akteurin nicht mehr habitualisiert, sondern entscheidet sich bewusst zu einer bestimmten Handlung angesichts einer Handlungsalternative.

Ein Ausweg aus diesem Dilemma besteht jedoch darin, den verantwortungsbewussten Umgang mit den eigenen Habitualisierungen als *wesentlich anlassbezogen* zu verstehen. Die Möglichkeit sich anders zu verhalten besteht nicht nur hypothetisch, sondern drängt sich zuweilen auch als bewusste Möglichkeit auf. Störungen in den eingeübten Praktiken bringen habitualisierte Einstellungen zu Bewusstsein und stellen Gründe dar, diese zu überdenken. Sie bilden

361 Wir sehen hier von der Möglichkeit ab, dass man auch einer Person den Vorwurf machen kann, bestimmte Erfahrungen überhaupt nicht gesammelt zu haben. Ich will nicht bestreiten, dass dies relevant ist in Hinblick auf problematische und ungerechtfertigte Formen von Vertrauen und Misstrauen. Jemand der grundlegend misstrauisch ist, und deshalb kaum überhaupt Vertrauensbeziehungen mit entsprechenden Erfahrungen eingeht, wird entsprechend auch schwieriger von seinem Vorurteil abzubringen sein. An dieser Stelle will ich aber ein anderes Problem adressieren, um die Idee der kontrafaktischen Sensitivität zu motivieren, die in der philosophischen Vertrauensliteratur eine zentrale Rolle spielt.

anlassbezogene Möglichkeiten Abstand zu den eigenen automatisierten Handlungsweisen zu gewinnen und diese zu überdenken. Vor dem Hintergrund dieser Idee muss der verantwortliche Umgang mit dem eigenen automatisierten Handeln nicht im Modus permanenter Reflexion gedacht werden – wodurch dieses nämlich gerade kein automatisiertes Handeln mehr darstellen würde. Explizite Reflexion ist vielmehr nur dann angemessen, wenn die konkreten Umstände Anlass dazu geben.

Man kann hier zwischen zwei verschiedenen Arten von Anlässen unterscheiden, eine vertrauende Einstellung zu überdenken. Eine erste einfache Kategorie bilden konkrete Erfahrungen von Vertrauensenttäuschung, die zumindest *ex post* eine Vertrauensbeziehung irritieren und Anlass geben können diese zu überdenken. Dies kann man im oben erläuterten Sinne als die Erfahrung der Störung einer Praxis verstehen. Vertrauenspraktiken haben stets einen Zweck, der, wie wir oben erläuterten entweder instrumentell mit den Vorzügen von gelingender Kooperation zusammenhängt, oder intrinsisch in der Vertrauensbeziehung als solcher liegt. In beiden Fällen aber kann dieser Zweck nur verwirklicht werden, wenn sich der Vertrauenspartner auch faktisch als vertrauenswürdig erweist. Vertrauensbrüche lassen sich somit auch als Störungen einer Vertrauenspraxis wenden: In diesen zeigt sich, dass die Praxis in der vorliegenden Form dem mit ihr verbundenen Zweck nicht erfüllen kann. Solche Störungen können und sollten somit ein Anlass sein, die Vertrauenspraxis ins Bewusstsein zu bringen, zu hinterfragen und entsprechend der gemachten Erfahrungen abzuändern.

Daneben gibt es aber natürlich auch Anlässe in *vorausschauender* Weise Vertrauen in andere zu überdenken, und nicht erst dann, wenn das Vertrauen bereits enttäuscht wurde. An diesem Punkt nun gewinnt die zuweilen in der Vertrauensliteratur geäußerte Idee an Attraktivität, dass sich Vertrauen wesentlich dadurch als vernünftig gestalten lässt, dass die Vertrauende eine Art kontrafaktische Sensitivität (*counterfactual sensitivity*) ausübt.[362] Dieses Konzept spielt im Vertrauensdiskurs insbesondere eine bedeutende Rolle im Zusammenhang mit Positionen, die wir oben bereits einmal als kontextualistische Positionen bezeichnet haben. Diese gehen davon aus, dass Vertrauen als eine *default*-Einstellung *prima facie* gerechtfertigt ist, solange es keine konkreten Gründe für Zweifel gibt. Auch diese Position wird zuweilen mit der Idee verknüpft, dass Vertrauende nicht permanent explizit über mögliche Gründe oder Gegengründe nachdenken müssen. Die Idee ist vielmehr, dass Vertrauende sich insofern rational verhalten können, als sie

362Der Begriff wurde meines Wissens maßgeblich von Elizabeth Fricker geprägt (1994) geprägt. Die dahinter stehende Idee als solche spielt aber vor allen Dingen für nicht-reduktionistische Positionen in der Zeugenschaftsdebatte eine zentrale Rolle. Siehe: Burge 1995; McDowell 1998, 430.

kontrafaktisch sensitiv gegenüber möglichen Anlässen für Zweifel eingestellt sind. Kontrafaktische Sensitivität bezeichnet demnach eine Einstellung der prinzipiellen Wachsamkeit gegenüber möglichen Gründen für Zweifel an den habitualisierten Vertrauenshaltungen. In der Definition von Elizabeth Fricker bedeutet es für eine Vertrauende kontrafaktisch sensitiv zu sein, dass „if there were signs of untrustworthiness, she would register them, and respond appropriately."[363]

Die Idee einer solchen latenten kritischen Disposition kann somit einen Mittelweg gegenüber dem oben skizzierten Dilemma markieren. Zum einen nämlich wird diese gerade nicht durch ständige explizite Reflexion ausgeübt, wodurch Vertrauen gerade seinen Charakter als *habitualisiertes* Vertrauen nicht verliert. Zum anderen wird durch den Umstand, dass die unterschwellige Sensitivität animiert werden kann, wenn ein konkreter Anlass Gründe für Zweifel liefert, Vertrauen auch in seiner vorbewussten Form als mehr oder weniger rational verständlich. Um ein mögliches Missverständnis zu vermeiden, sei noch angemerkt, dass es hier, in den Begriffen von Sanford Goldberg und David Henderson einen Unterschied gibt zwischen einem Wachsamsein (*being on the lookout*) gegenüber Anlässen für Zweifel und einem aktivem Suchen (*going out and looking*) nach solchen.[364] Mit dem Begriff der kontrafaktischen Sensitivität ist hier natürlich nur das *implizite* Wachsamsein gemeint. Das aktive Suchen nach Gründen für Zweifel wäre schließlich nicht mit einer genuin vertrauenden Haltung vereinbar. Insofern man nämlich nach Gründen für Zweifel sucht, vertraut man gerade nicht mehr.

Zu diesen Überlegungen kann noch ergänzt werden, dass ein entscheidender Anlass für Zweifel am Vertrauen in eine Person nicht nur epistemische Gründe, sondern auch schlichtweg die Erhöhung des Risikos darstellen kann, wobei zur Erinnerung gesagt sei, dass das Risiko hier definiert wird als die Kombination aus der Wahrscheinlichkeit einer Vertrauensenttäuschung und der Höhe des dadurch erlittenen Schadens.[365] Man betrachte etwa den folgenden Fall: Angenommen Person P erledigt alltäglich viele kleine Einkäufe über online-Transaktionen. In der Regel nimmt er diese routiniert vor und denkt nicht großartig darüber nach, wie sicher diese Form des Bezahlens eigentlich sei. Er hat gegenüber den Mechanismen des online-Bezahlens also ein habitualisiertes Vertrauen ausgebildet. Eines Tages jedoch nimmt er eine größere Anschaffung vor, für die er einen ungewöhnlich hohen Geldbetrag überweisen soll. Obwohl er nun, im Gegensatz zu vorher, keine zusätzlichen Gründe dafür hat, die gegen die Verlässlichkeit von online-Überweisungen sprechen, beginnt P zu zögern und fragt sich: Ist online-Bezahlung eigentlich

363Fricker 1994, 150.
364Goldberg/Henderson 2006.
365Dieser Punkt kann etwa auch an der kontextualistischen Position von Prijic-Samarzija (2007) ergänzt werden.

wirklich eine so verlässliche Angelegenheit? In diesem Fall scheint es mit der Erhöhung des schieren möglichen Schadens einen nachvollziehbaren Anlass für P zu geben an der Verlässlichkeit des online-Bezahlsystems zu zweifeln. Man mag dies so ausdrücken, dass mit dem gestiegenem möglichen praktischen Schaden schlichtweg auch die epistemischen *Standards* für vernünftiges Vertrauen gestiegen sind. Denn angesichts einer gestiegenen Schadenshöhe ist es vernünftiger, sich seiner Sache sicherer zu sein.

Prinzipiell gilt also: Angesichts eines gestiegenen, und sich damit vom Vertrauten abhebenden, Risikos gibt es einen vernünftigen *Anlass* dazu, über die Verlässlichkeit des Mechanismus bzw. die Vertrauenswürdigkeit der anderen Person nachzudenken und Gründe, die für oder gegen diese sprechen, abzuwägen. Erhöhtes Risiko kann aber nicht nur an der erhöhten Eintrittswahrscheinlichkeit des Schadens, und damit etwa an Gründen für Zweifel an der Vertrauenswürdigkeit festgemacht werden. Vielmehr kann auch die Erhöhung des Schadens ein entscheidender Anlass zu ausdrücklichen Überlegungen über die Vertrauenswürdigkeit der Vertrauensperson sein. Der Anlass für Modifikation oder Überdenken des habitualisierten Vertrauens ist also primär eine Erhöhung des Risikos, und dies umschließt auch eine Erhöhung der epistemischen Unsicherheit.

Dies sei aber nur in Parenthese angemerkt. Die entscheidende Einsicht war bisher, dass mithilfe der Idee der kontrafaktischen Sensitivität verständlich gemacht werden kann, inwiefern die Habitualisierung von Vertrauen zuweilen unthematisch, aber dennoch als Prozesse ablaufen kann, in deren Kontext sich Akteure mehr oder weniger rational verhalten. Der hier vorgeschlagene Ansatz erklärt dies über die Möglichkeit der Neujustierung habitualisierter Vertrauenseinstellungen im Lichte von Enttäuschungserfahrungen oder konkreter Anlässe für Zweifel. Dieser Ansatz erklärt also insgesamt inwiefern sich Vertrauende auch im Kontext und Umgang mit Habitualisierungsprozessen mehr oder weniger rational verhalten können, ohne dass dies darauf hinausläuft, dass sie diese permanent ausdrücklich reflektieren müssen. Im gleichen Moment erklärt dies, weshalb Akteure als verantwortliche Akteure beschrieben werden können, denen auch unreflektierte Vertrauens- oder Misstrauenseinstellungen zum Vorwurf gemacht werden können. Sie sind weiterhin verantwortlich, insofern sie sich zu diesen Habitualisierungen immer auch verhalten, und zwar mehr oder weniger kritisch und wachsam gegenüber Anlässen die Vertrauenseinstellungen zu überdenken. Und insofern Akteure darin scheitern, kann ihnen auch ein Vorwurf gemacht werden.

8.4 Vertrauen als wesentlich habitualisiert

Zu zeigen: kontext- und zeitübergreifend stabile Haltung

Mit den vorangegangen Überlegungen haben wir einen grundlegenden Kategorienwechsel in der Beschreibung von Vertrauen vorgenommen. Im bisherigen Verlauf der Arbeit haben wir wesentlich beschrieben, was Vertrauen als *Haltung überhaupt* ausmacht. Nun fokussieren wir darüber hinaus einen bestimmten Modus von Vertrauen, nämlich Vertrauen in Form der Kategorie einer auf wiederholten Erfahrungen gewachsenen habitualisierten *Praxis,* die nur aufgrund konkreter Anlässe hinterfragt wird. Dieser Kategorienwechsel leitet aber nicht einfach nur eine erweiterte Perspektive auf das Vertrauensphänomen ein, von der ausgehend sich einige neue interessante Aspekte des Vertrauensphänomens beleuchten ließen. Vielmehr scheint der im Alltag verankerte Vertrauensbegriff im Kern gerade auf Vertrauen in dieser habitualisierten Vollzugsform abzuzielen. Der Modus habitualisierter Praxis scheint also einen weiteren wesentlichen Aspekt des Vertrauensbegriffes zum Ausdruck zu bringen.

Motivieren wir diesen Punkt kurz mit einem Beispiel, das helfen wird unser vortheoretisches Sprachverständnis aufzudecken. Stellen wir uns einmal eine kleine Firma vor, die von einem sozial ungeschickten Chef geleitet wird, der auch die Buchhaltung übernimmt. Dieser Chef sei von einem eigentlich vollkommen pflichtbewussten und vertrauenswürdigen Charakter. Die Art und Weise wie er mit seinen Mitarbeitenden kommuniziert, lässt dies jedoch permanent infrage stellen und sorgt aufs Neue für Irritationen. Angenommen er macht permanent trockene Bemerkungen darüber, dass er vorhabe Bilanzfälschungen oder Steuerbetrug zu begehen. Die Mitarbeitenden sind dadurch permanent irritiert und sich nicht ganz sicher, ob diese Bemerkungen als trocken ironische Kommentare oder ernsthafte Ankündigungen aufzufassen sind. Nehmen wir weiterhin an, die Mitarbeitenden ringen aus diesen Gründen stets mit sich, wenn es zur Frage kommt, ob sie ihrem Chef voll und ganz mit der Buchhaltung vertrauen können. Sie mögen dann vielleicht stets zu sich sagen: „Das kann er nicht ernst sondern nur ironisch gemeint haben. Im Grunde genommen hat er sich ja immer vertrauenswürdig verhalten“. Sie müssen sich also stets dazu durchkämpfen ihm tatsächlich zu vertrauen.

In diesem Szenario wird man nun zwar nicht sagen wollen, dass die Mitarbeitenden ihrem Chef durch und durch misstrauen. Man mag aber ebenso geneigt sein zu sagen, dass hier *nicht wirkliches* Vertrauen gegenüber dem Chef herrscht. Die Tatsache, dass sich die Mitarbeitenden stets aufs Neue dazu bringen müssen ihrem Chef zu vertrauen, ohne dass sich so etwas wie eine stabile Praxis des Vertrauens ihm gegenüber ausgebildet hat, verleitet dazu sagen zu wollen, dass

hier kein volles Vertrauen ihm gegenüber besteht. Dies bringt zum Ausdruck, dass wir eingewöhntes Vertrauen implizit als das eigentliche Paradigma von Vertrauen verstehen. Der Inbegriff von Vertrauen zeigt sich hier im habitualisierten Vertrauen, das sich auf auf Grundlage wiederholter positiver Erfahrungen eingestellt hat. Eine Tendenz die der möglichen Habitualisierung entgegenläuft, wird deshalb als Ausdruck mangelnden Vertrauens gewertet.

Betrachten wir noch einen zweiten Fall, in dem es um Vertrauen um eine individuelle habitualisierte Praxis geht: Angenommen ich habe eine ausgeprägte Flugangst und kann mich niemals ruhig in ein Flugzeug setzen. Permanent kommen mir Gedanken darüber auf, was im schlimmsten Fall passieren kann. Womöglich aber mag ich mich stets kurzzeitig mit dem Gedanken beruhigen können, wie unwahrscheinlich ein Absturzszenario sei. Dennoch drängen sich mir immer wieder aufs Neue Zweifel auf. Wie wären wir in einem solchen Fall geneigt meine Einstellung zu beschreiben? Als vertrauend oder als zweifelnd? Einerseits mag man geneigt sein zu sagen, dass ich hier in keinem Fall einfach nur misstrauisch bin. Ich scheine vielmehr permanent zwischen einer vertrauenden und einer zweifelnden Haltung hin und her zu schwanken. Es drängen sich mir zwar temporär Bedenken auf, allerdings scheine ich diese im Einzelnen auch stets überwinden zu können: Ich kann mir selbst zureden, wie unwahrscheinlich ein Absturz ist und wieder zu Vertrauen gelangen oder schlichtweg die Bedenken vergessen. Andererseits mag es womöglich auch eine Neigung geben schlichtweg zu sagen, dass ich eigentlich *nie wirklich* vertraue. Der Umstand, dass sich mir wiederholt Zweifel aufdrängen und ich mich meines Vertrauens vergewissern muss, drängt einen dazu zu sagen, dass hier keine *wirklich* vertrauende Haltung vorliegt.[366] Auch hier zeigt sich also, dass der alltagssprachlich verankerte Vertrauensbegriff sein Paradigma in Form einer Vertrauenspraxis sieht, die solange unthematisch bleibt, als keine konkreten Störungen auftauchen. Eine Person die stetig und ohne Anlass über ihre Gründe für Vertrauen nachdenken muss, würden wir aus diesem Grund nicht als *wirklich* vertrauend beschreiben. Habitualisierung ist deshalb gerade der paradigmatische Modus der Ausbildung von Vertrauen.

Es ist zu beachten, dass sich Vertrauen in habitualisierter Form zwar als ein prototypischer Kern unseres Vertrauensverständnisses zeigt. Dies bedeutet aber natürlich nicht, dass damit Vertrauen in überlegter Form oder im Modus der Entscheidung *gar nicht* mehr als Vertrauen anzusehen ist. Es handelt sich immer noch um Vertrauen, das jedoch gegenüber dem *abfällt*, was sich in unserer Redepraxis als *eigentlicher* Kern von Vertrauen darstellt. Insofern diese

366Karen Jones (2012a, 240f.) äußert eine ähnliche Intuition, auch wenn sie diesen auf den Begriff des Selbstvertrauens bezieht.

Vertrauensformen also *überhaupt* am eigentlichen Wesenskern des Vertrauensbegriffes gemessen werden, mit diesem kommensurabel sind, bestätigt sich also gerade, dass es um Vertrauen geht – wenn auch um gewissermaßen „defizientes“ Vertrauen.

In der Einleitung dieses Kapitels nannten wir bereits einige Autor:innen, die verstreut postulieren, dass Vertrauen in seiner eigentlichen Form unreflexiv vorliegt und durch Reflexion gar unterminiert werde. Wir bezeichneten dies als die Hypothese der Reflexionslosigkeit von Vertrauen. Die Idee, für die wir soeben argumentiert haben, dass Vertrauen dem Wesen nach in habitualisierter Form vorliegt, kann nun eine Interpretationsfolie sein, durch die sich diese Annahmen in der Literatur in Bezug auf Vertrauen nachvollziehen lassen.[367] So ermöglicht sie beispielsweise eine bestimmte Lesart von Überlegungen von Oli Lagerspetz. Lagerspetz betrachtet den Fall, in dem eine Person einen Gast über Nacht in der Küche schlafen lässt, ohne auch nur ernsthaft darüber nachzudenken, ob dieser sie im Schlafe mit einem Küchenmesser erstechen könnte. Eben dieser Umstand, dass sie nicht gar nicht erst darüber nachdenkt, zeichne dabei das Vertrauen zum Gast aus: „a certain lack of reflexion and explicitness is *constitutive* of the trustfulness of [this] relation.“[368] Nicht über das Vertrauen nachzudenken, sei ein wesentliches Merkmal des Vertrauens zum Gast: „My trust in a guest shows itself in the fact that I did not think of my trust“[369]

Damit nimmt auch Lagerspetz eine Gestalt von Vertrauen in den Blick, die sich dadurch auszeichnet, dass die Vertrauenden ihr Vertrauen nicht ins Bewusstsein bringen, geschweige denn über mögliche Gründe für Zweifel an der Vertrauenswürdigkeit des anderen nachdenken. Wohlgemerkt scheint auch Lagerspetz damit nicht auf den oben genannten trivialen Punkt hinzuweisen, dass Vertrauen und Zweifel sich aus begrifflichen Gründen ausschließen, dass man also, *insofern* man vertraut nicht zugleich zweifeln kann. Dieser Punkt nämlich schließt nicht aus, dass Zweifel überwunden und man somit zu Vertrauen erst nach ausdrücklichem Überlegen *gelangen* kann. Lagerspetz aber trifft eine stärkere Feststellung. Er stellt heraus, dass die wiederholte und anlasslose Reflexion des Vertrauens bereits Ausdruck einer *nicht wirklich* vertrauenden Haltung ist. Dies betrifft, wie gesagt, nicht nur die Reflexion der Gründe für das Vertrauen. Auch eine Neigung dazu, sich das Vertrauen überhaupt bewusst zu machen, läuft einer tatsächlich vertrauenden Haltung gegen den Strich: „if we are very conscious of the fact that we trust one another, or keep talking about it, one will have some reason to wonder if there is really

367Darunter fällt womöglich auch das Luhman‘sche Theorem, dass Vertrauen funktional gesehen einen *entlastenden* Charakter hat.
368Lagerspetz 1998, 21.
369Ebd., 25.

trust between us at all."[370] Hier, so lässt sich Lagerspetz auslegen, verweist er ebenso wie wir in den obigen Beispielen auf die Idee, dass Vertrauen eigentlich in habitualisierter Form vorliegt.

Dies bedeutet zudem auch, dass wenn sich einem Vertrauenden Gründe aufdrängen, das Vertrauen in eine Person zu überdenken, damit bereits irreversibel der *Charakter* des Vertrauens in die Person verändert ist. Dies erläutert Lagerspetz an der folgenden Abwandlung des obigen Beispiels: Nehmen wir an, ich vertraue generell fremden Gästen, dass sie mich nachts nicht erstechen werden. Dann jedoch fragt mich jemand, wie ich denn Fremden so einfach vertrauen könne. Nehmen wir außerdem an, er berichtete von Fällen, in denen Fremde ihren Gastgeber nachts ermordet hätten. In diesem Fall, so Lagerspetz, kann es zwar sein, dass ich solche Bedenken letztendlich als nicht relevant einstufe. Man kann sich etwa vorstellen, dass ich mich darauf besinne, wie selten solche Fälle letztlich seien. Damit würde ich die Frage, ob ich dem Gast tatsächlich vertraue, letztlich weiterhin bejahen. Allerdings habe sich dennoch irreversibel etwas am Charakter meines Vertrauens gegenüber dem Fremden verändert: „we see graphically how the existence of a genuine question – no matter what the answer – already gives the situation a particular twist. If I recognise your question as a serious one, my perception of my guest will have changed-by that very fact."[371] Auch dieser Punkt lässt sich so deuten, dass Lagerspetz deutlich macht, dass das Vertrauen durch Bewusstwerdung irreversibel seinen präreflexiven Selbstverständlichkeitscharakter als habitualisierte Praxis verliert. Auch wenn ich mich nämlich letztendlich darauf besinne weiterhin zu vertrauen, hätte mein Vertrauen nachhaltig seine Ausprägungsform geändert.

Man mag gegenüber dieser Interpretation einwenden, dass Lagerspetz womöglich eigentlich einen anderen Punkt im Blick hat. So scheint Lagerspetz auch den besonderen Stellenwert der *Relation* zwischen den beiden Vertrauenspartner:innen zum Verständnis des Begriffes interpersonellen Vertrauens hervorheben zu wollen. Tatsächlich ist es exegetisch nicht von der Hand zu weisen, dass Lagerspetz auch darum bemüht ist, diesen relationalen Charakter von Vertrauen zu verdeutlichen. Allerdings muss dies nicht notwendigerweise seinem Bestreben widersprechen Vertrauen ebenso als im Wesentlichen habitualisiert auszuweisen. Ganz im Gegenteil lässt sich argumentieren, dass die Habitualisierung von Vertrauen einer Person gegenüber, oder eine Tendenz dies nicht zu tun, *gerade auch* Weisen darstellen, sich gegenüber der anderen Person zu verhalten. Auf dieses Thema werden wir unten im vierten Teil zurückkommen, wenn es um die Verknüpfung zwischen der epistemische und der anerkennungstheoretischen Seite

370Ebd., 29.
371Ebd., 21.

von Vertrauen geht.

8.5 Zwischenfazit

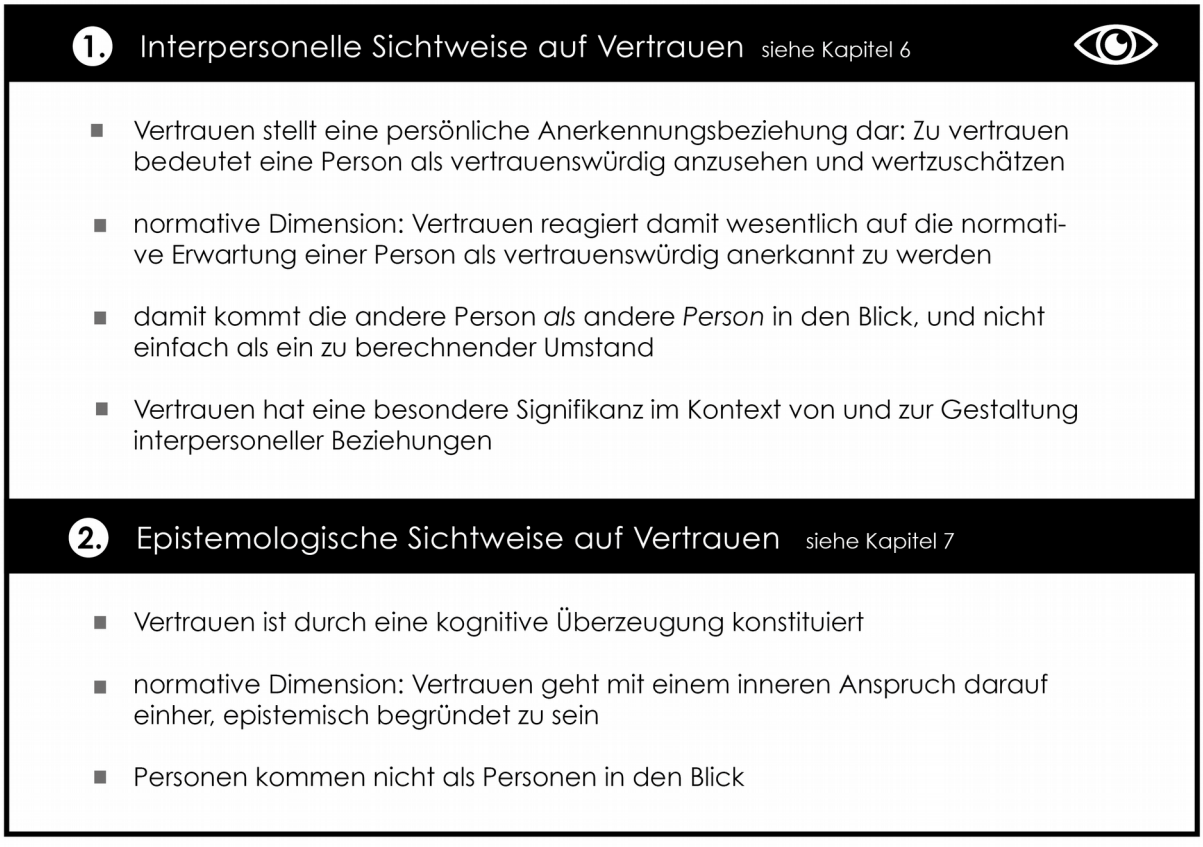
In diesem dritten Teil der Arbeit haben wir eine ausführliche Analyse der epistemischen Struktur von Vertrauen vorgenommen. Im Zentrum stand dabei der verworrene Zusammenhang zwischen Vertrauen und seinen epistemischen Gründen. Wir konnten grundlegend nachweisen, dass Vertrauen mit epistemischen Gründen in einem *begrifflichen* Zusammenhang steht, insofern Vertrauen eine kognitive Erwartung beinhaltet und mit einem inneren Anspruch einhergeht epistemisch begründet zu sein. Diesen Zusammenhang galt es aber in zwei Punkten zu präzisieren, die auseinanderzuhalten sind. Erstens bedeutet zu vertrauen in einem trivialen Sinne nicht über die Gründe für das Vertrauen zu reflektieren. Insofern man über die Gründe für Vertrauen nachdenkt, vertraut man nämlich gerade nicht. Davon ist zweitens der Punkt zu unterscheiden, dass Vertrauen zuweilen durch ausdrückliches Nachdenken untergraben oder verunmöglicht wird. Mit diesem Punkt haben wir uns zuletzt ausgiebig beschäftigt. Wir haben ihn so gedeutet, dass hier Vertrauen offenbar unter dem Gesichtspunkt der Kategorie einer habitualisierten Praxis in den Blick genommen wird. Diese verliert gerade durch Befragen nach ihrer Begründung ihren habitualisierten, quasi-natürlichen Charakter.

IV. Teil: Die Beziehung zwischen interpersoneller und epistemischer Seite

9. Integration der epistemischen und der interpersonellen Seite

9.1 Die Dichotomie zweier Sichtweisen auf Vertrauen

Bereits im Rückblick auf diese Arbeit mag einem auffallen, dass wir im zweiten und dritten Teil aus je sehr unterschiedlichen Perspektiven über Vertrauen gesprochen haben. Im zweiten Teil fokussierten wir auf Vertrauen als einem interpersonellen Phänomen, das eine normative Dimension hat und in der Anerkennung der anderen Person bestand. Im dritten Teil betrachteten wir Vertrauen in Hinblick auf seine kognitive Komponente und dessen epistemische Rechtfertigung, wobei sich diese Diskussion auf Vertrauen im generischen Sinne bezog, von dem interpersonelles Vertrauen eine Unterkategorie darstellt. Damit sind zwei Perspektiven auf das Phänomen interpersonellen Vertrauens aufgemacht, die bereits auf den ersten Blick bemerkenswert unterschiedlich ausfallen:



Graphik 3: Die interpersonelle und die epistemologische Betrachtung von Vertrauen.

Es stellt sich damit die Frage, wie diese beiden Sichtweisen sinnvoll miteinander zu vereinbaren

sind, wie die anerkennungstheoretische Seite interpersonellen Vertrauens mit seiner epistemischen Struktur zusammenhängt. Wie kann Vertrauen als epistemisch begründet gedacht werden, ohne dass es seinen Charakter als wesentlich interpersonelles Phänomen verliert? Wie können anders gesagt die praktische und die theoretische Seite von Vertrauen miteinander versöhnt werden?

In der Vertrauensliteratur wird der Gegensatz zwischen diesen beiden Beschreibungsperspektiven zuweilen in Form eines handfesten zentralen Dilemmas verhandelt. Dort werden die beiden unterschiedlichen Sichtweisen üblicherweise mit den Konzepten der *evidential view* und der *assurance view* belegt.[372] Insofern diese beiden als sich gegenseitig ausschließend verstanden werden, steht man vor einem Paradox, da beide Sichtweisen zugleich etwas Wesentliches über Vertrauen auf den Begriff bringen und damit notwendig erscheinen. Genauer gesagt lassen sich typischerweise folgende zwei Äste des Dilemmas rekonstruieren: Zum einen stellt sich das Problem, dass die interpersonelle Sichtweise auf Vertrauen an sich noch nicht deutlich macht, inwiefern Vertrauen epistemisch gerechtfertigt gedacht werden kann. Diese Seite des Dilemmas hat beispielsweise Jennifer Lackey in ihrer Kritik an der uns bekannten *assurance view* von Moran deutlich gemacht.[373] Wenn das Vertrauen in die Zeugenschaft einer Person sich nämlich allein darauf stütze, dass der Sprecher eine Versicherung abgibt, bliebe es epistemisch vollkommen willkürlich. Immerhin, so lautet in aller Kürze Lackeys Argument, versichert auch ein Lügner oder eine schlecht informierte Person die vermeintliche Wahrheit ihrer Aussage. Solange also nicht gesagt werden kann, ob der Sprecher auch *vertrauenswürdig* ist, bleibt dessen Versicherung ohne epistemische Relevanz. Die bloße *assurance view* liefert damit keine epistemische Basis für das Vertrauen in den Sprecher und droht dieses also als blind ausweisen zu müssen. Dieses Problem zeichnet sich besonders scharf ab, wenn man bedenkt, dass man zwei Personen A und B gegenüberstehen kann, die Gegenteiliges behaupten.[374] In diesem Fall wäre man mit zwei Versicherungen konfrontiert, ohne, dass man beiden zugleich Vertrauen schenken könnte. Vielmehr muss man sich entweder einer Beurteilung enthalten oder sich dazu entscheiden entweder A oder B zu vertrauen. Sofern man sich für das Vertrauen in eine der beiden Personen, sagen wir für A, entscheidet, stellt sich die Frage nach den Gründen für diese Wahl. Und diese Gründe können nicht allein in der Tatsache bestehen, dass A eine Versicherung gegeben hat, denn dies trifft ebenso auf B zu. Es braucht also Gründe für das Vertrauen in A, die über die Tatsache der bloßen

372 Vgl. Faulkner 2007, 2011; Lackey 2008, insbes. Kap.6; Simpson 2018.

373 Wobei sich Lackey in ihrer Kritik nicht exklusiv auf Moran bezieht, sondern eine allgemeinere Klasse verschiedener Ansätze kritisiert, die sie unter den Begriff des *Interpersonal View of Testimony* fasst. Für unsere Zwecke werden wir den Kern ihrer Kritik anhand von Morans Ansatz erläutern.

374 Diese Idee zur Zuspitzung der Problembeschreibung übernehme ich aus: Kaminski 2020, 101.

Versicherung von A hinausgehen.

Den anderen Ast des Dilemmas stellt die von Moran selber formulierte Kritik an der *evidential view* dar, nach der es diese verfehlt auf den Begriff zu bringen, was es eigentlich heißt einer Person zu glauben, und nicht einfach nur zu glauben, was die Person sagt. Wenn Vertrauen nämlich als Ergebnis eines bloßen evidenzbasierten Kalküls gedacht würde, hätten wir es gerade nicht mehr mit veritablem interpersonellem Vertrauen zu tun. Wir vertrauen dann, wie wir oben mithilfe von Hertzberg formuliert haben, nicht mehr der anderen *Person*, sondern bestenfalls den Gründen, die wir haben um davon auszugehen, dass die Person vertrauenswürdig ist. Damit schiene man hier gewissermaßen auf die *evidential view* zurückzufallen. Somit offenbart sich hier insgesamt ein Dilemma, dass sich für Lackey zwar unter einem bestimmten Vorzeichen darstellt, das sie aber in einer Formulierung auf den Punkt bringt, die auch für unsere Situation ganz passend ist: Vertrauen ist demnach entweder "genuinely interpersonal but epistemologically impotent, or it is not epistemologically impotent but neither is it genuinely interpersonal."[375]

Wir sind zwar oben mit unserer Anerkennungstheorie des Vertrauens von der *assurance view* abgewichen und im Gegensatz zu Lackey nicht allein am Problem der Zeugenschaft interessiert. Dennoch stellt sich für uns in ganz ähnlicher Weise die Frage nach der Versöhnung der interpersonellen und der epistemischen Perspektive auf Vertrauen. Immerhin bleibt mit der Anerkennungstheorie, wie wir gesehen haben, die Fragen nach der epistemischen Rechtfertigung von Vertrauen zunächst offen. Anerkennung haben wir schließlich als das Einsehen eines normativen Anspruches der Vertrauensperson verstanden, das innerhalb einer Beziehung zur anderen Person situiert ist und zugleich eine Rolle zur Gestaltung dieser Beziehungen spielt. In dieser Beschreibung scheint es aber ein rein interpersonell-normatives und kein epistemisches Phänomen darzustellen. Anders ausgedrückt: In der Perspektive des anerkennungstheoretischen Blickes auf Vertrauen, wird zwar deutlich, inwiefern Vertrauen einem Gründe *liefern* kann. Wenn ich gefragt werde, woher ich weiß, dass p der Fall ist, kann ich als Grund ausführen: „Weil A es mir gesagt hat und ich ihm vertraue". Offen bleibt jedoch die Frage, auf welchen Gründen das Vertrauen *selber* wiederum beruht. Und hier kann man nicht wieder auf das Vertrauen als solches selber verweisen ohne sich in einen Begründungszirkel zu begeben. In dieser anerkennungstheoretischen Hinsicht ist also noch keine epistemische Rechtfertigungsbasis für Vertrauen ersichtlich.

Die Gefahr dieser fehlenden epistemischen Absicherung von Vertrauen stellt dabei für

375Lackey 2008, 222.

Lackey im Besonderen deshalb ein Problem dar, da es ihr darum zu tun ist zu begründen, wie man *Wissen* auf Grundlage der Bezeugungen anderer gewinnen kann. Das Wissen gründet sich immerhin auf dem Vertrauen in andere Personen. Wenn dieses Vertrauen jedoch wiederum vollkommen epistemisch arbiträr wäre, bleibt unverständlich, wie wir auf einer solchen schwachen Basis gerechtfertigte Überzeugungen gewinnen können. Auf der Grundlage unserer obigen Überlegungen zum inneren epistemischen Anspruch von Vertrauen können wir dieses Problem auch vom Zeugenschaftskontext lösen und allgemeiner darstellen. Wenn nämlich für Vertrauen im Allgemeinen gilt, dass es intrinsisch mit dem Anspruch verbunden ist epistemisch angemessen zu sein, ergibt sich auch für dieses, dass es nicht zugleich als epistemisch vollkommen willkürlich verstanden werden kann. Somit zeigt sich auch in Bezug auf unsere bisherigen konzeptionellen Weichenstellungen und in Bezug auf Vertrauen im Allgemeinen (losgelöst vom Vertrauen in Zeugenschaft) das strukturell gleiche Problem, das Lackey ansprach: Beide Seiten auf Vertrauen, die interpersonelle und die epistemische, scheinen notwendig, stehen aber dichotom nebeneinander. Wie sind sie also zusammen zu denken?

In diesem letzten Kapitel soll es darum gehen diese Dichotomie zwischen der interpersonellen und der epistemischen Seite von Vertrauen aufzuheben. In der Literatur lassen sich insbesondere zwei auf den ersten Blick vielversprechende Strategien dazu identifizieren. Eine Gruppe von Ansätzen zielen darauf ab, die Brücke zwischen der interpersonellen und der epistemischen Seite von Vertrauen auf additive Weise zu schlagen. Ich nenne sie deshalb additive Lösungsansätze. Demnach trifft die Beschreibung der interpersonellen Dimension von Vertrauen zu, sie müsse lediglich noch durch epistemische Überlegungen *ergänzt* werden. Das Problem an solchen Ansätzen jedoch zeigt sich darin, dass der epistemische Aspekt von Vertrauen nicht vollständig von der interpersonellen Seite getrennt betrachtet werden kann. Dies offenbaren insbesondere Fälle, in denen Vertrauen zwar epistemisch abgesichert ist, aber auf eine Art und Weise, dass sie dem Vertrauen in die andere Person eigentlich gerade entgegenlaufen. Ohne genauer qualifizierte Vorstellung zwischen den beiden Seiten von Vertrauen, schließt der additive Ansatz aber solche Fälle nicht aus, geschweige denn, dass er eine Antwort darauf liefert, was genau an solchen Fällen problematisch ist (9.1.1). Ein zweiter Ansatz, der sich auf die sogenannte *entitlement*-Position im Zeugenschaftsdiskurs bezieht, versucht demgegenüber das Verhältnis von Vertrauen und epistemischen Gründen so zu denken, dass solche problematischen Fälle scheinbar ausgeschlossen sind. Wie wir zeigen werden, stellt dies aber nicht wirklich eine globale Lösung des eigentlichen Problems dar (9.1.2). Unser eigener Vorschlag wird demgegenüber mit einer erneuten

Reflexion des Begriffes der personalen Anerkennung beginnen und nachweisen, inwiefern dieser *intrinsisch* mit epistemischen Gründen verbunden ist (9.1.3). Die Anerkennung einer Person und der Bezug zu epistemischen Gründen stehen demnach gerade nicht im Widerspruch zueinander. Ganz im Gegenteil, epistemische Gründe spielen sogar eine konstitutive Rolle zur Anerkennung des Vertrauenspartners. Im Anschluss daran stellt sich jedoch noch eine Folgefrage, da die genauen Zusammenhänge weiterhin an einem Punkt dunkel bleiben: Die Vertrauensbeziehungen, in denen wir uns befinden, stellen auch einen Hintergrund dar, vor dem wir Evidenzen für oder gegen Vertrauen bewerten. Dadurch droht Vertrauen aber zu einer sich selbst stützenden Einstellung und also irrational zu werden. Zu diesem möglichen Missverständnis werden wir einige abschließende Bemerkungen liefern (9.2).

9.1.1 Der additive Lösungsansatz

In der Literatur gibt es eine oft und in verschiedene Versionen vertretene Antwort auf die Herausforderung die interpersonelle und die epistemische Seite von Vertrauen zu integrieren, die wir als Varianten des *additiven Lösungsansatz* bezeichnen können. Unter diesem verstehen wir hier eine ganz grundlegende Struktur einer bestimmten, verschiedentlich ausbuchstabierten, Antwortrichtung.[376] Die prinzipielle Idee besteht dabei darin, zunächst an der interpersonellen Beschreibung von Vertrauen – in unserem Fall wäre dies die Beschreibung von Vertrauen als einer anerkennenden Haltung – als der korrekten Beschreibung dessen, was Vertrauen im Kern ausmacht, festzuhalten. Auf der anderen Seite wird lediglich *ergänzend* darauf hingewiesen, dass dieses Vertrauen zudem auch epistemisch begründet sein sollte. Die interpersonelle und die epistemische Seite von Vertrauen sind damit aber nebeneinander gestellt, ohne, dass eine interne Beziehung zwischen den beiden postuliert oder erläutert wird. Diesen Weg deutet etwa Richard Moran an einer Stelle seines Aufsatzes an. So hat Moran den oben beschriebenen ersten Ast des Dilemmas zwischen den beiden Beschreibungsperspektiven wohl im Blick und gesteht ein: „[…] the idea is not that the speaker‘s authority over the constitution of the particular speech act he is

376Neben Moran (2005, 16.) findet sich diese Struktur auch bei Benjamin McMyler (2011, Kap. 5.3/5.4). McMyler unterscheidet strukturell zwischen: erstens der Beschreibung der interpersonellen Situation zwischen den beiden Vertrauenspartnern, die er in einem bestimmten Sinne als zweitpersonal aufeinander eingestellt versteht. Und zweitens der epistemischen Annahme des Vertrauenden, dass bestimmte Hintergrundbedingungen erfüllt sind, die das Vertrauen auch situationsspezifisch rechtfertigen. Auch bei ihm stehen damit aber die interpersonelle Beschreibung und die epistemische Rechtfertigung unerläutert nebeneinander. Für ihn stellt sich damit zumindest die Frage, wie er die in diesem Kapitel beschriebenenen Fälle defizienten Vertrauens als solche ausschließen will.

performing (e.g., as assertion rather than recitation) shoulders the epistemic burden all by itself.“[377] In Reaktion darauf hält er fest, dass die *assurance view* zwar weiterhin eine treffliche Beschreibung der Beziehung zwischen Sprecher und Hörer liefere. Damit der Hörer jedoch dem Sprecher gerechtfertigterweise glauben könne, müsse er *auch* eine *zusätzliche* epistemische Bewertung (*epistemic assessment*) der Vertrauenswürdigkeit des Sprechers vornehmen.

In ähnlicher Weise deutet sich diese methodische Trennung bei Paul Faulkner an. Faulkner geht zunächst von der bekannten Moran‘schen *assurance-view* aus, nach der der Vertrauenspartner (hier der Sprecher) seinen Sprechakt nicht als Präsentation eines Beweise für das Gesagt, sondern vielmehr als eine Versicherung versteht, die darauf abzielt anerkannt zu werden. Dem Sprecher zu vertrauen heißt wiederum, diese Versicherung zu akzeptieren.

Über dieses uns bereits bekannte Theorem hinaus betont Faulkner jedoch, dass das Vertrauen selber dem Vertrauenden einen spezifischen Grund dafür gibt, anzunehmen, dass sich der Vertrauenspartner auch vertrauenswürdig verhalten wird.[378] Faulkner geht davon aus, dass zu vertrauen bedeutet, davon auszugehen, dass derjenige dem vertraut wird, in der Abhängigkeit des Vertrauenden einen Grund dafür sieht, das in ihn gesetzte Vertrauen nicht zu enttäuschen. Entsprechend besagt die bereits einmal zitierte Vertrauensdefinition von Faulkner: „A trusts S to Φ (in the affective sense) if and only if (1) A depends on S Φ-ing; and (2) A expects (1) to motivate S to Φ (where A expects this in the sense that A expects it of S that S be moved by the reason to Φ given by (1)).“[379]

Ein Grund für die Annahme des Vertrauenden, dass sich der Vertrauenspartner vertrauenswürdig verhalten wird, liegt damit im Vertrauen selber, weshalb Faulkner festhält: „the act of trust is rationally self-supporting.“[380] Hinter dieser Idee haben einige Kritiker:innen[381] jedoch eine viziöse „bootstrapping“-Struktur vermutet. Es sieht so aus, als könne sich Vertrauen selber epistemisch rechtfertigen und in Münchhausen‘scher Manier eigenhändig aus dem Sumpf des epistemisch Willkürlichen ziehen. Damit jedoch wäre Vertrauen immer und „automatisch“ epistemisch vernünftig. Man könne dann nicht mehr erklären, warum wir manchen Personen berechtigterweise vertrauen und anderen nicht. Tatsächlich: Selbst wenn wir annehmen, dass der Vertrauensnehmer einen Grund dafür haben mag, sich vertrauenswürdig zu verhalten, es handelt sich dabei nur um einen *pro tanto* Grund, in Bezug auf den nicht klar ist, dass er sich in den

377Moran 2005, 16.
378Damit grenzt sich Faulkner von den üblichen nicht-reduktionistischen Theorien ab, die davon ausgehen, dass Vertrauen durch keinerlei positiven Grund gerechtfertigt sei.
379Faulkner 2011, 146.
380Ebd.,151.
381So etwa: Hinchman 2012; Longworth 2012.

praktischen Überlegungen des Vertrauensnehmers durchsetzt. Entscheidend ist schließlich nicht, welche Gründe der Vertrauenspartner prinzipiell erwägt, sondern welchen er „all things considered*"*, ein derartiges *Gewicht zumisst*, dass er *aus* ihnen handelt. Zu wissen, dass jemand vertrauenswürdig ist, bedeutet demnach auch mehr, als zu wissen, welche Gründe er überhaupt in Erwägung zieht. Es bedeutet vielmehr, zu wissen, welche Gründe letztlich für ihn motivational ausschlaggebend sind.[382]

Insbesondere als Reaktion auf seine Kritiker:innen hat Faulkner deshalb präzisiert, dass es zwar stimmt, dass Vertrauen nur eine Art *pro tanto*-Grund für den Vertrauenspartner liefert, um sich vertrauenswürdig zu verhalten. Dies stelle eine adäquate Beschreibung der interpersonellen Situation zwischen den beiden Vertrauenspartnern dar. Jedoch sei dies nicht hinreichend, um zu erklären, wann genau Vertrauen epistemisch angemessen sei. Dafür wiederum brauche es *zusätzlich* weitergehende epistemische Gründe, die sich auf die Situation beziehen, in der sich der Vertrauende befindet[383]: „So while it is true that an actual explanatory connection between trust and trustworthiness could not ground the *initiation* of trust, what grounds this initiation is not this connection but how the subject thinks about the trust situation in adopting the attitude of trust (with the presumptions it expresses).“[384] Was genau die epistemischen Gründe angeht, stellen sich nun ganz spezifische Probleme bezüglich der Position Faulkners, auf die ich hier nicht im Detail eingehen werde.[385] In jedem Fall gilt: Um eine angemessene Antwort auf seine Kritiker:innen zu geben, müssen diese epistemischen Gründe für die Initiation von Vertrauen begründen können, warum wir einer konkreten Person A vertrauen und einer anderen Person B misstrauen, sprich weshalb wir eine als ausreichend vertrauenswürdig ansehen, eine andere jedoch nicht. Die beschriebenen *pro-tanto* Gründe selber können solche Unterscheidungen nicht rationalisieren, denn sie sind überhaupt nur ausschlaggebend, wenn man es bereits mit einer Person zu tun hat, die eben vertrauenswürdig ist. Es braucht also von diesen *pro-tanto* Gründen unabhängige Gründe für die Annahme der Vertrauenswürdigkeit einer konkreten anderen Person.[386] Ein solcher Ansatz zur

382Dieses Problem artikuliert etwa auch Annette Baier: „Trusting is rational, then, in the absence of any reason to suspect in the trusted strong and operative motives which conflict with the demands of trustworthiness as the truster sees them.“ (Baier 1986, 24).

383Vgl. Hinchman 2012, 48.

384Faulkner 2012, 118.

385Die entscheidenden Gründe zur epistemischen Absicherung von Sprechervertrauen sieht Faulkner dabei in der Etablierung sozialer Normen dazu, die Wahrheit zu sagen. Damit aber bleibt die Frage nach der Vertrauenswürdigkeit der *konkreten Person* vor mir (siehe die Kritik von Hinchman (2012, 55)). Wenn ich einer Person A vertraue und Person B misstraue, dann weil ich die Überzeugung habe, dass A vertrauenswürdig ist und, dass B nicht vertrauenswürdig ist. Ich brauche also irgendwelchen epistemischen Anhaltspunkte für diese spezifisch *personenbezogenen* Überzeugungen.

386Zugegeben: Faulkner selbst sperrt sich zuweilen gegen die Idee, dass Vertrauen auf solchen personenbezogenen epistemischen Gründen beruht, da er dazu neigt, dies mit der *evidential view* gleichzusetzen (vgl. Faulkner 2007).

Integration der beiden Seiten von Vertrauen wäre damit aber auch bloß additiv strukturiert. Er liefert zum einen eine Beschreibung der Situation zwischen den beiden Vertrauenspartnern, und stellt dieser zum anderen schlichtweg zur Seite, dass das Vertrauen auch als epistemisch angemessen begründet sein muss. Damit stehen die beiden Seiten ohne Erläuterung eines inneren Zusammenhanges nebeneinander.

Historisch erinnern die beiden additiven Vorschläge von Moran und im Anschluss an Faulkner im Übrigen an eine in der *Logik von Port Royal* vertretene Auffassung darüber, wie aus Zeugenschaftsberichten gewonnene Überzeugungen gerechtfertigt werden können.[387] Hier findet sich eine ähnliche hybride Struktur vor, indem die Autoren zum einen darauf abzielen, eine adäquate (nicht-reduktionistische) *Beschreibung* dessen vornehmen, was sie eine theoretische Autorität nennen. Dabei gilt, dass i*nsofern* jemand eine theoretische Autorität darstellt, ihm in seinem Urteil zu glauben ist. Davon getrennt besteht jedoch die Frage, *wer* konkret *als* theoretische Autorität eigentlich anerkannt werden sollte. Dies bedarf nämlich einer epistemischen Bewertung der potenziellen Autorität.

Die additiven Ansätze haben also insgesamt gemeinsam, dass sie zum einen versuchen eine adäquate Beschreibung der interpersonellen Beziehung zwischen den beiden Vertrauenspartnerinnen vorzunehmen und zum anderen lediglich *ergänzend* darauf hinweisen, dass das Vertrauen, sofern es nicht einfach als arbiträr gedacht werden soll, auf epistemischen Gründen beruhen muss, die dafür sprechen, dass der Vertrauenspartner auch vertrauenswürdig ist. Diese Ansätze zielen auf der einen Seite also darauf ab, eine phänomenangemessene Beschreibung vom Vertrauen in eine Person zu liefern, während sie auf der anderen Seite anmerken, dass dieses Vertrauen epistemisch gerechtfertigt sein muss. Beide Seiten werden dabei wohlgemerkt konzeptionell auseinandergehalten. Ein analytischer Bezug zwischen beiden wird nicht deutlich gemacht, etwa indem man ausgehend von einer Betrachtung der interpersonellen Seite von Vertrauen als solcher bereits dessen notwendigen Zusammenhang zu epistemischen Gründen aufzeigte. Es wird lediglich darauf hingewiesen, dass es *beide* Seiten zum angemessenen Verständnis von Vertrauen brauche.

Im Folgenden werden wir nicht die verschiedenen Varianten des additiven Lösungsansatzes im Einzelnen diskutieren. Stattdessen soll grundlegend die geteilte additive Struktur dieser Vorgehensweisen problematisiert werden. Das zentrale Problem derselben ist, dass sie nicht genau

Sein eigener Versuch zur Umgehung des Dilemmas zwischen der *evidential view* und der *assurance view*, indem er Vertrauen als eine Art bloßer Annahme versteht, verstrickt sich jedoch in ein Dilemma. Siehe hierzu: Kaminksi 2020, 155-179.

387Siehe die Darstellung in: McMyler 2011, Kap. 1.

spezifizieren, wie das Vertrauen in eine Person mit epistemischen Gründen zusammenhängt. Dadurch lassen sie durchaus problematische Fälle zu, in denen die Art, wie eine Person auf epistemische Gründe Bezug nimmt, gerade einem Vertrauen in die andere Person widerspricht. Schauen wir uns dafür ein extremes Beispiel an: Angenommen ein Ladenbesitzer möchte sicher gehen, dass seine Angestellt:innen kein Geld aus der Ladenkasse entwenden. Zu diesem Zweck installiert er heimlich im gesamten Laden Überwachungskameras, um die Mitarbeitenden auszuspionieren.[388] In diesem Falle hätte er eine ausreichende epistemische Grundlage, um davon auszugehen, dass seine Angestellt:innen vertrauenswürdig sind. Allerdings scheint er in der spezifischen *Art*, wie er die epistemische Evaluierung derselben vorgenommen hat, gerade kein Vertrauen, im Gegenteil sogar im höchsten Grade Misstrauen zum Ausdruck zu bringen. Auch wenn er also eine epistemische Basis zur Bestätigung der Vertrauenswürdigkeit der Angestellt:innen hätte, stellt dies gerade eine Basis dar, auf der es begrifflich unmöglich wird zu sagen, dass er ihnen *vertraut*. Hier scheint wieder ein Fall vorzuliegen, für den wir sagen können, dass eine Person eigentlich nur ihren Gründen vertraut, die dafür sprechen, dass die andere Person vertrauenswürdig ist, ohne der anderen Person zu vertrauen.

Das Beispiel zeigt, dass sich genuines Vertrauen in eine Person grundlegend mit bestimmten Formen der epistemischen Absicherung ausschließt. Insbesondere bestimmte Formen der Überwachung einer anderen Person oder des Suchens nach Gründen für deren Vertrauenswürdigkeit scheinen gerade im Widerspruch zu einer vertrauenden Haltung dieser gegenüber zu stehen. Paul Faulkner selber merkt an einer Stelle gar diesen Punkt an und hält deshalb fest, dass Vertrauen grundlegend eine „anti-monitoring attitude“[389] darstelle. An den entscheidenden Stellen bleibt aber unklar, inwiefern sich sein Ansatz von den genannten problematische Fällen abgrenzt. Und auch Morans Begriff der epistemischen Bewertung (*epistemic assessment*) bleibt nicht klar von solchen problematischen Formen der Überwachung abgegrenzt. Der Weg der additiven Ansätze ist also nicht so einfach gangbar, da mit diesen nicht klar ausgeschlossen ist, das problematische Formen der epistemischen Bewertung der Vertrauenswürdigkeit einer Person ausgeschlossen sind. Nennen wir dies kurz das *Problem defizienter oder pathologischer Vertrauensfälle.*

Die Fälle zeigen also, dass bestimmte Weisen sich auf epistemische Gründe zu beziehen, gerade einem Vertrauen in die Person widersprechen. Deshalb legt sich nahe, dass sich aus einem

388Der Fall ist angelehnt an ein Beispiel von Arnon Keren (2014, 2607), der damit auf einen ähnlichen Punkt hinzuweisen beabsichtigt. Siehe auch: Faulkner 2007, 882.
389Faulkner 2011, 87.

genaueren Verständnis dessen, was es heißt einer Person zu vertrauen, auch eine ganz bestimmte Charakterisierung des Zusammenhanges zwischen interpersonellen Vertrauen und epistemischen Gründen ergeben muss. Die Aufgabe wird es also sein, diesen inneren Zusammenhang zwischen Vertrauen und Gründen zu eruieren, um vor diesem Hintergrund klarer zu erkennen, was genau die genannten defizienten Vertrauensfälle problematisch macht.

Bevor wir darauf eingehen, soll im Folgenden aber noch eine Position diskutiert werden, die verspricht einen alternativen Mittelweg aufzuzeigen. Diese stellt scheinbar eine Möglichkeit dar, die epistemische Rationalität von Vertrauen so zu denken, dass Fälle defizienten Vertrauens *a priori* ausgeschlossen sind. Dafür bezieht sie sich auf das sogenannte Konzept der *entitlement*. Diese recht technische Diskussion ist jedoch für den Argumentationsverlauf nicht notwendig. Die Leserin kann diesen deshalb auch überspringen.

9.1.2 Exkurs: Zum Lösungspotenzial von *Entitlement*-Ansätzen

In der Zeugenschaftsdiskussion spielen solche Ansätze eine prominente Rolle, die davon ausgehen, dass es eine globale Rechtfertigung dafür gibt den Aussagen anderer zu vertrauen. Diese besagen, dass es *im Normalfall* gerechtfertigt ist den Aussagen anderer zu vertrauen, das heißt zumindest solange kein konkreter Anlass dafür spricht an der Aussage zu zweifeln. Der Kern dieser Positionen wird in der Regel in einer Annahme zusammengefasst, die Jonathan E. Adler als die *default rule of testimony* bezeichnet. Diese besagt Folgendes: „If the speaker S asserts that p to the hearer H, then, under normal conditions, it is *correct* for *H* to *accept* (believe) *S*'s assertion, unless *H* has special reasons to object."[390] Solange also keine Gründe vorliegen, die dagegen sprechen einer Person zu vertrauen, ist das Vertrauen in die andere Person bereits ausreichend begründet.

Diese Sichtweise kennt eine ganze Reihe von Variationen, die die Diskussionslage auf diesem Gebiet einigermaßen unübersichtlich machen. Eine Vielzahl an möglichen Taxonomien der Positionsvarianten auf diesem Feld wäre denkbar. An dieser Stelle werden wir aber nur grob zwei Arten von Ansätzen unterscheiden. Ihr Unterschied macht sich daran fest, was eigentlich als eine Standard-*Rechtfertigung* des Vertrauens verstanden wird. Hier gibt es zum einen Ansätze, die die Rechtfertigung im Sinne eines allgemein geltenden epistemischen *Grundes* auffassen.[391] Vertrauen

390Siehe auch die ähnlichen frühen Formulierungen dieses Prinzips in: Price 1969, 5. Vorlesung.
391Verschiedene Versionen dieser Position unterschieden sich in der Regel anhand der Erklärung, die sie für die

in die Zeugenschaft anderer sei demnach deshalb im Allgemeinen gerechtfertigt, da es einen positiven Grund gibt, der dafür spricht, dass Sprecher in der Regel vertrauenswürdig sind. Wie gesagt kann sich im Einzelfall stets herausstellen, dass es konkrete Gegengründe gibt, die gegen die Vertrauenswürdigkeit eines bestimmten Sprechers sprechen, sodass das Vertrauen in ihn alles in allem doch nicht gerechtfertigt sein kann. Die besagt Position postuliert also nur, dass es einen *pro tanto*-Grund dafür gibt der Zeugenschaft anderer zu vertrauen. Eine davon abweichende Sichtweise besagt, dass die Standardrechtfertigung für Vertrauen im Sinne einer sogenannten „Berechtigung" (*entitlement)* zu verstehen sei.[392] Dies stellt eine Form von Rechtfertigung dar, die gerade auf keinerlei positiven Gründen basiert. Wir werden darauf im Folgenden genauer eingehen.

Den Hintergrund dieses Konzeptes bildet eine Unterscheidung von Tyler Burge zwischen zwei verschiedenen Arten von Rechtfertigungen (*warrants*) einer epistemischen Überzeugung. Im Folgenden werde ich für diese die englischen Originalbegriffe verwenden: *justifications* und *entitlements*. Für Burge sind mit *justifications* dem Erkenntnissubjekt zugängliche und von ihm angebbare Schlussfolgerungen auf eine bestimmte epistemische Überzeugung gemeint: Eine *justification* „must be available in the cognitive repertoire of the subject."[393] *Justifications* stellen also dem Erkenntnissubjekt zugängliche *gründebasierte* Überlegungen dar, die auf eine bestimmte Überzeugung schließen lassen.[394] *Entitlements* beziehen sich demgegenüber unmittelbar auf eine zuverlässige Erkenntnisquelle. Über diesem Wege bedarf es keiner *Schlussfolgerung*, um zu einer Überzeugung zu gelangen. Vielmehr ist das Erkenntnissubjekt aufgrund der allgemeinen Verlässlichkeit der Erkenntnisquellen unmittelbar dazu berechtigt eine Überzeugung aus dieser zu beziehen.[395] Diese Art der Rechtfertigung spielt vor allen Dingen für die oben bereits erwähnten non-reduktionistischen Positionen eine entscheidende Rolle, die Zeugenschaft als eine eigenständige Erkenntnisquelle, neben etwa Perzeptionen, Inferenzen oder Erinnerung, verstehen wollen. Üblicherweise wird es als unmittelbares Korollar dieser These angesehen, dass der Zeugenschaft anderer unter normalen Bedingungen Glauben geschenkt werden kann.[396]

Annahme eines solchen Grundes geben. Für einen Überblick siehe etwa: Adler 2006, Abschnitt 2.

392Siehe: Burge 1995.

393Ebd., 273.

394Unklar ist bei Burge allerdings, ob sich *justifications* ausschließlich als internalistische und *warrants* als bloß extternalistische Rechtfertigung verstehen lassen. Burge selber scheint dies zwar zuweilen so zu sehen (vgl. Burge 2003, 273). Allerdings bleiben dabei noch entscheidende Unklarheiten offen. Zur Diskussion dieses Punktes, siehe: Majors 2015.

395"But entitled, we do not have to have reasons that support the default position, if there is no reasonable ground for doubt. Truth telling a norm that can be reasonably presumed in the absence of reasons to attribute violations. (ebd., 281)

396Eine Ausnahme bildet die Theorie von Paul Faulkner (2011), der beansprucht eine non-reduktionistische Position zu vertreten, aber zugleich davon ausgeht, dass es positive Gründe dafür brauche, um einer Person zu vertrauen.

Erläutern wir, was unter einem *entitlement* gemeint ist, kurz anhand des Beispiels der Erinnerung. Wenn mir meine Erinnerung sagt, dass ich heute morgen einen Kaffee getrunken habe, dann weiß ich dies, da ich mich *direkt* auf meine Erinnerung verlassen kann. Die Rechtfertigung dafür ebendies zu tun liegt darin, dass meine Erinnerung schlichtweg eine verlässliche Erkenntnisquelle ist. Dabei brauche ich also nicht von der *Prämisse*, dass mein Erinnerungsvermögen verlässlich ist, auf die Einsicht zu *schließen*, dass ich heute morgen einen Kaffee getrunken habe: „In fact [I] needn't reason from premises at all in this case: [I] can be warranted in his memory belief simply by trusting his memory – as long as [my] memory is reliable on such a matter."[397] Um zu wissen, dass ich einen Kaffee getrunken habe, muss ich mich nur daran erinnern dies getan zu haben. Erst wenn ich etwa feststelle, dass ich jeden Morgen so schläfrig bin, dass sich zuweilen meine vorangegangenen Träume, mit dem, was wirklich passiert in meiner Erinnerung vermischen, dann liegen im konkreten Fall Umstände vor, die dagegen sprechen, dass ich mich direkt auf mein Erinnerungsvermögen verlasse.

Gemäß der non-reduktionistischen Position kann nun auch Zeugenschaft als eine ebensolche verlässliche Erkenntnisquelle betrachtet werden. Die Rechtfertigung des Vertrauens in Zeugenschaft erklärt sich dann, ebenso wie im Falle des Gedächtnis, nicht in Form eines inferentiellen Schlusses, es handelt sich also nicht um eine Rechtfertigung im Sinne einer *justification*. Vielmehr gibt es eine unmittelbare *entitlement* dafür, den Mitteilungen anderer zu vertrauen, sofern keine Indizien für deren mangelnde Vertrauenswürdigkeit vorliegen. Entsprechend formuliert Burge das folgende Prinzip namens Akzeptanz-Prinzip (*Acceptance Principle)*, das eine spezifische Formulierung der oben dargestellten *default rule of testimony* darstellt: „A person is entitled to accept as true something that is presented as true and that is intelligible to him, unless there are stronger reasons not to do so."[398]

Als Vertrauender vertraut man damit gemäß dieses Gedankens nicht notwendigerweise insofern rational, als man alle erdenkbaren Gründe für und wider abwägt, sondern als man sich in dem übt, was Tyler Burge und an ihn anknüpfend Edward Hinchman als kontrafaktische Sensitivität („counterfactual sensitivity") bezeichnet – ein Konzept, das wir bereits aus der obigen Diskussion zur Rationalität habitualisierter Vertrauenspraktiken kennen. Hinchman hält entsprechend fest: „[T]rust is reasonable insofar as it is governed by a counterfactual sensitivity to evidence of untrustworthiness: if there is, or if there had been, evidence that S is not worthy of

397Hinchman 2005, 578. Vgl. Burge 1995, 276.

398Burge 1995, 281. Wir beschäftigen uns hier nicht mit der genauen Herleitung dieses Prinzips von Burge. Zur Kritik daran, siehe: Faulkner 2011, 93-112.

your trust, you will cease trusting S, or you would not have trusted S in the first place."[399] Kontrafaktische Sensitivität beschreibt damit eine Aufmerksamkeit in Bezug auf mögliche negative Gründe, die gegen die Vertrauenswürdigkeit der anderen Person sprechen und eine Bereitschaftshaltung diese in Erwägung zu ziehen.

Losgelöst von den Debatten um auf Zeugenschaft beruhendem Wissen interessiert und hier nur die Konzeption von *entitlement* und *justification* als Modell, um über die epistemische Rechtfertigung von Vertrauen im Allgemeinen nachzudenken. Demnach ließe sich die epistemische Rechtfertigung von Vertrauen im Sinne von *entitlement* und nicht im Sinne von *justification* auffassen. Dann ist Vertrauen nicht insofern gerechtfertigt, als man positive Gründe für die Annahme hat, dass die andere Person vertrauenswürdig ist. Vielmehr ist es *im Allgemeinen* gerechtfertigt, solange *keine* besonderen Gründe *dagegen* sprechen.

Natürlich ließe sich in einer ersten Problematisierung die grundlegende Annahme des Ansatzes als solches hinterfragen. Ist es tatsächlich angemessen von einer solchen Form von *entitlement* im Bezug auf Vertrauen überhaupt auszugehen? Hier ließe sich etwa die Kritik von Elizabeth Fricker oder Paul Faulkner im Bezug auf Zeugenschaftsvertrauen aufgreifen und auf Vertrauen als solches verallgemeinern. Beide liefern Argumente, nach denen nichts dafür, sondern eher dagegen spreche, dass andere Personen *im Normalfall* als vertrauenswürdig anzusehen sind. Wäre dem so, dann bräuchte man letztendlich doch wieder *positive Gründe* für diese Annahme. Auf diese Diskussion müssen wir uns in Hinblick auf unsere eigentlich Fragestellung jedoch nicht einlassen. Im Folgenden werden wir nämlich zeigen, dass selbst wenn *a fortiori* der *entitlement*-Vorschlag plausibel ist, dieser als solcher noch nicht weiterhilft das Problem der defizienten Vertrauensfälle zu umgehen.

Auf den ersten Blick scheint die *entitlement*-Sichtweise einen Lösungsraum zu begründen, innerhalb dessen die Dichotomie zwischen der praktischen und der theoretischen Sichtweise auf Vertrauen aufgehoben werden kann, und dabei das Problem defizienter Vertrauensfälle umgeht.[400]

399Hinchman 2014, 24.

400Weiterhin gibt es auch einen Ansatz von Thomas Simpson (2018), der allerdings keinen *entitlement*-Ansatz zu verfolgen scheint, sondern versucht zu begründen, dass es *pro tanto*-Gründe für das Vertrauen in andere Sprecher gibt. Ganz ähnlich wie der im Folgenden diskutierte *entitlemen*-Ansatz von Hinchman, scheint auch Ansatz Simpsons nur auf den ersten Blick für unser hier diskutiertes Problem einschlägig. Simpson geht es primär darum zu zeigen, dass es im Normalfall epistemisch gerechtfertigt ist einem Sprecher zu glauben, und das der Grund dafür letztlich in den etablierten Normen der Zeugenschaftspraxis wurzelt. Demnach gelte: „it is normally rational for (third-personal) belief to follow on from (second-personal) acceptance of a changed normative relationship" (ebd., 8) Damit argumentiert er für einen Begründungszusammenhang nachdem aus der normativen Verpflichtung des Sprechers gegenüber dem Hörer für den Normalfall folge, dass er die Wahrheit sagt. Dieser Ansatz ist zunächst deshalb ungeeignet, um unser Problem zu adressieren, da unklar ist, wie verbreitet letztlich solche Normalfälle sind. Simpson stellt zwar fest, dass Sprecher im Groben und Ganzen vertrauenswürdig sein müssen, was zumindest heißt, dass nicht alle permanent lügen können. Dennoch lässt die Argumentation Simpsons zu, dass Menschen in

So wurde diese Idee auch in einem Aufsatz von Edward Hinchman ins Spiel gebracht, um eine konzeptionelle Integration der interpersonellen und der epistemischen Seite von Vertrauen zu vorzuschlagen. Hinchman sieht dabei die gründebasirete epistemische Bewertung der Vertrauenswürdigkeit der anderen Person zunächst als prinzipiell problematisch an. Dieses sei nicht mit einer Haltung personaler Anerkennung, und damit was es heißt *jemandem* zu vertrauen, vereinbar. Er unterstreicht „that assessing S [a speaker] for reliability is a way of refusing S's invitation to trust, since S is inviting A to trust *her*, not whatever evidence she thereby gives of her reliability."[401] Hinchman buchstabiert dieses Problem natürlich vor dem Hintergrund seines spezifischen zeugenschaftstheoretischen Ansatzes aus, deren Details wir an dieser Stelle aber nicht erläutern müssen. In der grundlegenden Struktur stellt sich das Problem bei Hinchman wie bei uns oben dar: Vertrauen steht demnach einerseits unter epistemischen Anforderungen, die er als „doaxastic-deliberative requirements" bezeichnet. Das Räsonnieren über solche Gründe aber steht für Hinchman im Widerspruch mit einer genuin vertrauenden Haltung, die er ebenfalls im bestimmten Sinne als anerkennende Haltung bezeichnet. Er akzentuiert deshalb in dem obigen Zitat, dass den epistemischen Gründe für die Vertrauenswürdigkeit einer Person zu vertrauen, gerade nicht bedeutet *der anderen Person* zu vertrauen.

Als Reaktion auf dieses Problem nun verweist Hinchman darauf, dass die epistemische Rechtfertigung von Vertrauen in Form eines *entitlement* konzipiert werden könne. Demnach müsse die epistemische Rechtfertigung nicht durch gründebasierte Reflexion, sondern kann auch durch kontrafaktische Sensitivität abgesichert werden: „A [a hearer] can fulfill his doxastic-deliberative requirements by exercising the counterfactual sensitivity informing his capacity for reasonable trust. If S is (or were) unreliable, or there is (or were) evidence that she is, then A will (or would) not reasonably trust. But if she is reliable, and there is no evidence that she is not, then A is entitled to believe that p by simply taking her word for it, without assessing evidence at all."[402]

Um das Problem der Fälle defizienten Vertrauens stimmig zu umgehen, scheint dieser Vorschlag aber nur bedingt hilfreich. Es stimmt zwar, dass mit ihm die epistemische Rechtfertigung

einem *Großteil* der Fälle weiterhin die Unwahrheit sagen. Der schwerwiegendere Einwand lautet aber, dass der *pro tanto*-Grund im konkreten Einzelfall nur dann ausschlaggebend sein kann, insofern keine konkreten gewichtigen Gründe für Zweifel vorliegen. An diesem Punkt bleibt unsere oben formulierte Frage unbeantwortet: Wie einerseits potenzielle Gründe für Zweifel relevant sein können, ohne dass man auf der anderen Seite wieder Formen von defizienten Vertrauens annehmen muss?

401Hinchman 2014, 17.

402Vgl. auch seine Darstellung in: Hinchman 2005. Hier bezieht sich Hinchman ebenfalls auf die das Konzept des *entitlements*, um damit eine Zwischenposition zwischen reduktionistischen und nicht-reduktionistischen Ansätzen zu formulieren: „The reductionist is right that it is reasonable to trust only those who are worthy of your trust. And the anti-reductionist is right that reasonable trust does not rest on an assessment of your interlocutor as worthy of it." (Hinchman 2005, 597)

von Vertrauen überhaupt nicht mehr im Sinne *einer epistemischen Abwägung* aller Gründe für und wider gedacht werden muss. Gerechtfertigtes Vertrauen muss nicht zwangsläufig so gedacht werden, dass positive epistemische Gründe dafür vorliegen, zumindest in bestimmten Kontexten reicht es, wenn keine negativen Gründe vorliegen, die gegen die Vertrauenswürdigkeit der anderen Person sprechen. So kann man laut *entitlement*-Ansatz epistemisch dazu berechtigt sein zu vertrauen, ohne in irgendeiner Weise darüber Gründe *abgewogen* zu haben. Zumindest in einem Fall solcher Art mag es dann so aussehen, als seien die beiden Sichtweisen auf Vertrauen so miteinander in Einklang zu bringen, dass sich das Problem der defizienten Vertrauensfälle nicht stellt: Das Vertrauen soll insofern als rational gelten, als man sensitiv gegenüber möglichen negativen Gründen ist. Damit ist ein Modell vorgelegt, in dem Vertrauen als epistemisch vernünftig vorgestellt werden kann und Fälle defizienten Vertrauens zugleich eindeutig ausgeschlossen scheinen. Der Fall des obigen Ladenbesitzers etwa stellt gemäß dem Hinchman‘schen Vorstoß keinen Fall von Vertrauen dar, da er sich nicht in kontrafaktischer Sensitivität übte, sondern aktiv epistemische Gründe für oder wider die Vertrauenswürdigkeit seiner Angestellten suchte und erwägte.

Zwei Einwände seien nun gegen diesen Vorschlag genannt. Zum einen hilft dieser Punkt bei der Versöhnung der beiden Perspektiven deshalb nicht weiter, da er nur einen *bestimmten Bereich* ausmacht, für den die Fälle defizienten Vertrauens ausgeschlossen sind. Nach dem Akzeptanz-Prinzip besteht die *entitlement* nämlich nur insoweit keine negativen Gründe vorliegen, die Zweifel an der Vertrauenswürdigkeit des Sprecher begründen würden.[403] Unabhängig von der Frage, wie verbreitet solche negativen Gründe tatsächlich sind – Elizabeth Fricker etwa geht, wie gesagt, davon aus, dass wir im Normalfall gerade Gründe dafür haben an der Vertrauenswürdigkeit von Sprechern zu zweifeln: „we know too much about human nature to want to trust anyone, let alone everyone, uncritically“[404] – es bleibt damit die Frage, wie Hinchman Fälle einordnen würde, in denen das Akzeptanz-Prinzip nicht greift, ohne, dass wir es in irgendeiner Form mit defizienten Fällen von Vertrauen zu tun hätten. In solchen Fällen nämlich, in denen die *entitlement* nicht angemessen ist, können wir unter unter Umständen dennoch einer Person vertrauen. Es gibt schließlich schlichtweg auch Fälle, in denen wir Gründe für die Vertrauenswürdigkeit einer Person abwägen, ohne, dass dies einem der defizienten Vertrauensfälle gleichkäme. Stellen wir uns etwa vor, dass ein Person A mir mitteilte, dass p der Fall sei, während mir Person B mitteilte, dass non-p der Fall sei. Es gibt für beide Mitteilungen als solche betrachtet keinen Anlass dazu, an der

403Zur Skepsis in Bezug auf die Reichweite des *Acceptance Principle*, siehe auch: Gelfert 2014, 75f.
404Fricker 1995, 400.

Vertrauenswürdigkeit des anderen zu zweifeln. Nehmen wir weiterhin an, es sei für mich von enormer Wichtigkeit zu wissen, ob p der Fall ist oder nicht. Vor diesem Hintergrund mag ich noch einmal in mich gehen und mich auf die möglichen Gründe besinnen, die für oder gegen das Vertrauen in A bzw. B sprechen könnten. Nehmen wir etwa an, A sei eine mir bekannte Person, von der ich schon öfter gehört habe, dass sie sich als vertrauenswürdig erwiesen hat. Ich entscheide mich also aufgrund dieser positiven Überlegung dafür ihr zu vertrauen. In diesem Fall nehme ich eine ausdrückliche gründebasierte Bewertung der Vertrauenswürdigkeit der Vertrauensperson vor, aber ohne, dass wir deswegen sagen würden, ich würde ihr eigentlich gar nicht vertrauen.

Es ist unklar, wie Hinchman solche Fälle einordnen würde. Wir haben es dabei schließlich nicht einfach mit einem Fall von *entitlement* zu tun. Die Sache ist komplexer, da auf jeden Fall positive Gründe und nicht einfach nur eine *pro tanto* Rechtfertigung für das Vertrauen eine Rolle spielen. Dennoch scheint es nicht problematisch solche Fälle als Fälle genuinen Vertrauens anzusehen. Es gibt einfach Vertrauensfälle, in denen die zugrundeliegende epistemische Reflexion komplexer sein mag, ohne, dass dies hieße, dass man eigentlich gar nicht der anderen Person vertrauen würde. Solche, vermutlich auch nicht seltenen Fälle, stehen jedoch außerhalb der Betrachtung Hinchmans. Würde er diese etwa als keine wirklichen Fälle von Vertrauen einstufen, da man sich dabei aktiv aufgrund positiver Gründe dafür entscheidet einer Person zu vertrauen? Damit bliebe der Ansatz entschieden revisionistisch. Denn was spräche dagegen hier davon zu sprechen, dass ich A vertraue? Insgesamt liefert Hinchman so gesehen keine globale Theorie zur Versöhnung der epistemischen und interpersonellen Seite von Vertrauen, die eine tragfähige Grenze zwischen allen denkbaren unproblematischen und defizienten Formen von Vertrauen ziehen könnte.

Der zweite Einwand gegen Hinchmans Vorstoß ist insofern tiefgreifender, als er auch die vermeintlich „unverdächtigen" Fällen, in denen das *entitlement*-Prinzip einschlägig ist, infrage stellt. Dies hängt mit der Frage zusammen, ob die Bezugnahme auf das Konzept der kontrafaktischen Sensitivität nicht doch wieder problematische Fälle zuließe. Das liegt darin, dass das Konzept theoretisch nicht besonders scharf bestimmt bleibt. Dadurch bleibt es offen, dieses etwa auch so wie Elizabeth Fricker aufzufassen, die einmal auf folgende Weise formulierte, was es heißt, kontrafaktische Sensitivität auszuüben: „the hearer must always be scrutinising the speaker for telltale signs of its absence, and she must be alert to the presence of such signs."[405] An dieser Stelle mag man sich auch wiederum fragen, ob eine solche Form der, wenn auch unbewussten,

405Fricker 1995, 405.

Wachsamkeit gegenüber Anzeichen, die die Vertrauenswürdigkeit der anderen Person infrage stellen, nicht doch wieder einem genuinem Vertrauen in die andere Person widerspricht. Es stellt sich hier zumindest die *Frage* danach, wie eine solche Wachsamkeit gegenüber Indizien für mangelnde Vertrauenswürdigkeit nicht doch wieder eine Form defizienten Vertrauens darstellt? In der Charaktersierung von Fricker sieht es so aus, als ob man in der Perspektive des *entitlement*-Ansatzes letztendlich auch nur der *kontrafaktischen Sensitivität*, aber nicht der anderen Person vertrauen würde. Insgesamt also hilft der *entitlement*-Ansatz nicht wirklich weiter eine geeignete und global gültige Integration der beiden Sichtweisen auf Vertrauen zu formulieren, die es vermag defiziente Vertrauensformen als solche auszuschließen.

9.1.3 Der innere Zusammenhang von Anerkennung und epistemischen Gründen

Der additive Lösungsansatz, so haben wir gesehen, lieferte keine überzeugende Zusammenführung der epistemischen und der interpersonellen Sichtweise auf Vertrauen, da er die internen begrifflichen Beziehungen zwischen beiden Dimensionen nicht aufzuzeigen vermag. Dies zeigte das Problem der defizienten Vertrauensformen auf. Der *entitlement*-Ansatz wiederum lieferte nur auf den ersten Blick eine Lösung hierzu, da er nur auf eine Klasse an Situationen abhebt, in denen – auch nur scheinbar – Fälle defizienten Vertrauens ausgeschlossen sind. Im Folgenden wird es deshalb darum gehen, aufzuzeigen, dass es einen intrinsischen analytischen Zusammenhang zwischen der für Vertrauen konstitutiven Form von Anerkennung einer Person und epistemischen Gründen gibt. Vor dem Hintergrund des Verständnisses dieses Zusammenhanges kann auch auf den Begriff gebracht werden, inwiefern Fälle defizienten Vertrauens diesem nicht gerecht werden. Auf dieser Grundlage kann zudem noch einmal grundlegend deutlich gemacht, was als das eigentliche Problem der *evidential view* anzusehen ist. Das Problematische an dieser besteht nämlich nicht in deren Annahme, dass es epistemische Gründe braucht, um einer Person zu vertrauen. Vielmehr liegt es darin, dass sie damit zusammenhängend Vertrauen zu einem bloß epistemischen Phänomen *reduziert*.

Anerkennung, so hatten wir es oben definiert, besteht darin einen normativen Anspruch einer anderen Person einzusehen, ihr somit also einen normativen Status zuzuschreiben. Im Falle von Vertrauen bezieht sich die Anerkennung auf die normative Erwartung des Vertrauenspartners als vertrauenswürdig angesehen zu werden. Dass nun ein intrinsischer Zusammenhang zwischen

Anerkennung und epistemischen Gründen besteht, kann in einem ersten Schritt motiviert werden, wenn man über die *Angemessenheitsbedingungen* dieser normativen Erwartung des Vertrauenspartners reflektiert. Diese orientiert sich nämlich immer schon an der epistemischen Ausgangslage.

Mit Anscombe hatten wir festgestellt, dass es manchmal eine Verletzung oder eine Beleidigung sein kann, wenn einem nicht vertraut wird. Zuweilen kann dies sogar auch ein Anlass für Empörung und Vorwürfe sein, etwa der Form „Warum vertraust du mir denn nicht?“ Wir hatten diesen Umstand so gedeutet, dass seitens des Vertrauenspartners ein Anerkennungsgesuch besteht: Er erwartet vom anderen als vertrauenswürdig angesehen zu werden. Was bei Anscombe jedoch nicht ausgeführt wird ist, dass sich auch prinzipiell etwas über die *Bedingungen* sagen lässt, unter denen Reaktionen persönlicher Verletzung oder Beschwerden *angemessener* oder weniger angemessen sein können. Dies hängt damit zusammen, dass es je nach situativen Umständen mehr oder weniger angemessen sein kann, vom anderen zu erwarten, dass er einen als vertrauenswürdig anerkennt. Extreme Beispiele bringen diesen Zusammenhang besonders klar zum Vorschein. Erinnern wir uns dafür etwa an das folgende von Anscombe beschriebene Szenario: Nehmen wir an, ein Lügners spreche nur deshalb die Wahrheit, da er falsch informiert sei. Der Hörer wiederum wisse, dass der Sprecher ihn täuschen will, dieser zugleich aber falsch informiert ist. Er schließe deshalb darauf, dass das, was der Sprecher sagt der Wahrheit entspricht, ohne, dass man sagen könnte, dass dieser jenem vertraut. Nehmen wir nun an, der Sprecher mit Täuschungsabsicht käme dahinter, dass ihm nur aufgrund eines Kalküls geglaubt würde – er also nicht als vertrauenswürdig anerkannt wurde. In diesem Fall würden wir dem Lügner wohl kaum einen berechtigten Anspruch dazu zugestehen, sich ob der mangelnden Anerkennung als vertrauenswürdig zu beschweren. Insofern der Sprecher eindeutig in keiner Weise vertrauenswürdig *ist*, ist er auch nicht dazu *berechtigt* zu erwarten, als vertrauenswürdig angesehen zu werden.

Schauen wir uns noch ein zweites Beispiel an: Stellen wir uns vor, eine unbekannte Person würde mich auf der Straße ansprechen und darum bitten, ihr eine exorbitant große Summe Geld zu überweisen, die sie für eine dringende Finanzinvestititon brauche. Sie verspricht mir jedoch, das Geld mit dem Gewinn, den sie daraus schlagen wird, inklusive eines kleinen Bonus an mich zurück zu zahlen. Nehmen wir weiterhin an, ich würde mich in diesem extrem kuriosen Fall nachvollziehbarerweise nicht auf das Versprechen der fremden Person einlassen, ihr also nicht vertrauen. Unter diesen konkreten Umständen scheint es, dass die andere Person kaum berechtigterweise in der Position wäre, mir mein mangelndes Vertrauen vorzuwerfen. Gleiches gilt

in Bezug auf emotionale Reaktionen, etwa persönliche Verletzung. So mag sich die fremde Person zwar persönlich durch mein Misstrauen verletzt fühlen. Es scheint jedoch fragwürdig, ob wir diese emotionale Reaktion tatsächlich als der Situation angemessen verstehen würden.[406]

Die zwei Fälle zeigen zwei Sachverhalte auf: Erstens ist die normative Erwartung des Vertrauenspartners als vertrauenswürdig anerkannt zu werden nur dann berechtigt, wenn er tatsächlich auch vertrauenswürdig *ist*. Zweitens ist sie nur angemessen, wenn angenommen werden kann, dass die potenziell vertrauende Person auch *ausreichende* epistemische Anhaltspunkte dafür hat, um von der Vertrauenswürdigkeit des Vertrauensnehmers ausgehen zu können.

Dieser zweite Punkt zeigt sich zudem auch anhand des Umstandes, dass sich aufgrund einer *Veränderung* der epistemischen Situation in der Regel auch verändert, was Personen an Vertrauen voneinander erwarten: Stellen wir uns etwa vor, dass innerhalb einer intimen Beziehung, wie einer Freundschaft, Person A gegenüber B einen schweren Vertrauensbruch begangen habe. Nehmen wir etwa an, A erzählt ein intimes Geheimnis von B weiter, das er eigentlich versprochen hatte, für sich zu behalten. Mit einer solchen Vertrauensenttäuschung ändert sich im Normalfall auch, was A gegenüber B in Zukunft angemessenerweise an Vertrauen erwarten kann. Wenn B nun nicht mehr so leichtwillig oder gar nicht mehr A vertraut, stellt dies angesichts des vorangegangen Vertrauensbruch keinen adäquaten Anlass für A dar, sich zu beschweren oder sich verletzt zu fühlen. Ein Vorwurf wie „Warum vertraust du mir nicht einfach?“ ist deshalb nicht mehr angemessen, da A mit seinem Vertrauensbruch seine Vertrauenswürdigkeit infrage gestellt hat. Wenn es offenkundige Gründe für Misstrauen gibt, kann Vertrauen auch nicht sinnvollerweise erwartet werden. Auch hier sieht man also, dass die epistemische Situation auch auf die normative Situation zwischen den beiden Interaktionspartnern einwirkt: Im Lichte der Erfahrung eines Vertrauensbruches kann A deshalb nicht mehr im gleichen Maße angemessenerweise erwarten, als vertrauenswürdig angesehen zu werden.

Solche Beispiele weisen also einen strukturellen Zusammenhang zwischen der normativen und der epistemischen Situation zwischen den beiden Vertrauenspartner nach. Demnach kann Vertrauen nur insofern erwartet werden, als davon ausgegangen werden kann, dass der Vertrauende auch ausreichende *epistemische* Anhaltspunkte dafür hat. Die normative Erwartung des Vertrauenspartners und die Anerkennung desselben durch den Vertrauenden sind somit immer *im Kontext epistemischer Gründe* situiert. Mit der Feststellung dieses qualitativen Zusammenhanges

406Dass Emotionen mehr oder weniger angemessen sein können, setzt übrigens ein gewisses Verständnis derselben voraus. Sie können demnach nicht als bloße Empfindungszustände, sondern müssen als kognitiv strukturiert aufgefasst werden. Siehe hierzu etwa die Emotionstheorie von Bennett Helm, etwa in: Helm 2011.

soll natürlich nicht ausgeschlossen sein, dass es im je konkreten Fall uneindeutig und umstritten sein mag, was genau angemessenerweise an Vertrauen erwartet werden kann. Hier mag es durchaus weniger klare Grenzfälle: Nehmen wir etwa an, der verzweifelte Fremde in dem Beispiel von oben, bitte mich nicht darum eine horrende Summe zu überweisen, sondern darum, mein Handy auszuleihen, da er einen wichtigen Anruf zu tätigen hat. Nehmen wir außerdem an, dies fände in einer Gegend statt, in der durch diese Masche bekannterweise häufig Handys gestohlen werden. Hat der Fremde nun einen berechtigten Anlass, darüber empört zu sein, wenn ich ihm misstraue? Über diese Frage mag man sich streiten. Der Umstand aber, dass es konkrete Fälle gibt, in denen es uneindeutig und kontrovers sein mag, was *genau* zu tun angemessen wäre, widerspricht natürlich nicht der These, dass sich die normative Erwartung, als vertrauenswürdig angesehen zu werden, *prinzipiell* an epistemischen Anhaltspunkten orientiert. Es mag lediglich im Einzelfall uneindeutig sein, welche epistemischen Anhaltspunkte genau bereits ausreichend zur Rechtfertigung einer solchen normativen Erwartung sind.

Um die Plausibilität der These, dass sich Anerkennung unmittelbar auf epistemische Gründe bezieht, weiterhin zu stützen kann man auch darauf aufmerksam machen, dass diese phänomenologisch adäquat einfängt, wann Vertrauen tatsächlich als anerkennend *empfunden* werden kann. Stellen wir uns dafür etwa eine Person vor, die in ihren Vertrauensentscheidungen vollkommen unvorhersehbar agiert, so als würde sie schlichtweg auswürfeln, wem sie ihr Vertrauen entgegenbringt und wem nicht. Unter diesen Voraussetzungen ist es zweifelhaft, ob man es tatsächlich als *anerkennend* empfinden würde, sollte man das Glück haben, dass einem diese Person nun einmal zufällig vertraut – immerhin hätte es *genauso gut* auch passieren können, dass sie einem misstraut! Was also an ihrem „Vertrauen“ wäre anerkennend? Man kann diesen Punkt auch so ausdrücken: Oben haben wir den anerkennenden Charakter von Vertrauen auch dadurch zum Ausdruck gebracht, dass wir Vertrauen als eine Form des Ernstnehmens des Vertrauensnehmers bezeichneten. Im Fall der radikal arbiträr „vertrauenden“ Person, würde diesem „Ernstnehmen“ jedoch gerade der *Ernsthaftigkeitscharakter* fehlen. Auch diese Überlegung unterstreicht also, dass gerade wenn wir Vertrauen als Anerkennung verstehen, sich diese intrinsisch auch auf epistemische Gründe beziehen muss. Gerade weil Vertrauen ein Anerkennen ist, kann man nicht arbiträr vertrauen. Gerade gemäß dem Anerkennungsverständnis ist „arbiträres Vertrauen“ also ein logisches Unding oder stellt maximal eine Form von Pseudo-Vertrauen dar.

Wir können nun vor dem Hintergrund dieser Überlegungen auch unsere obige Anerkennungsthese ein wenig präzisieren. Oben formulierten wir, dass vom Vertrauensnehmer ein

normativer Anspruch ausgeht als vertrauenswürdig angesehen zu werden. Wir sollten nun hinzufügen, dass der Anspruch des Vertrauensnehmers gegenüber dem Vertrauenden vielmehr darin besteht, dass dieser jenen *insofern* als vertrauenswürdig ansieht, als er *einsieht, dass epistemische Gründe dafür oder zumindest nicht dagegen sprechen, dass jener auch vertrauenswürdig ist.*

An diesem Punkt wird besonders deutlich, wie sich diese Überlegungen strukturell vom additiven Ansatz abgrenzen. Dieser ging lediglich davon aus, dass die interpersonell-praktische Beschreibung von Vertrauen durch eine epistemische Dimension *ergänzt* werden müsse. Damit werden die interpersonelle und die epistemische Seite von Vertrauen nur summiert, also weiterhin konzeptionell getrennt voneinander gedacht. Wir stellen demgegenüber heraus, dass es sich hier nicht einfach um eine Ergänzungsbeziehung handelt. Vielmehr weist ein genaueres Verständnis von Vertrauen als einer anerkennenden Haltung an sich bereits einen inneren Bezug zur epistemischen Situation auf. Die Analyse der obigen Beispiele zeigen, dass die für Vertrauen konstitutive Form der Anerkennung *intrinsisch* mit epistemischem Gründen in Beziehung steht.

Diese Analyse ist nun zudem vorbereitend, um im Folgenden das Problem der defizienten Vertrauensfälle anzugehen. Es stellt sich schließlich immer noch die Frage danach, inwiefern manche Formen sich auf epistemische Gründe zu beziehen, gerade nicht mit einer vertrauenden Haltung vereinbar sind. Denken wir etwa an das obige Beispiel des Ladenbesitzers, der seine Angestellten ausspionierte, um zu ermitteln, ob sie vertrauenswürdig sind. Solche Fälle hatten wir bereits oben in vortheoretischer Weise als im Widerspruch zu einer genuin vertrauenden Haltung markiert. Nun haben wir aber zumindest einen theoretischen Ansatzpunkt erarbeitet, um genauer auf den Begriff bringen, was im Kern das Problematische dieser Fälle ausmacht, sodass erklärt werden kann, weshalb sie allenfalls als pathologische Vertrauensformen anzusehen sind.

Ich möchte vorschlagen, dass der Schlüssel, um zu verstehen, was an den Fällen defizienten Vertrauens problematisch ist, darin liegt, diese Fälle *von der normativen Erwartung des Vertrauensnehmers her zu denken*. Demnach bestimmt sich was als vertrauensangemessene Form der Absicherung durch Gründe (und was nicht) angesehen werden kann dadurch, *was genau* der Vertrauenspartner in einer konkreten Situation angemessenerweise vom Vertrauenden erwartet. *Welche* epistemischen Gründe *genau*, die für seine Vertrauenswürdigkeit sprechen, verlangt er vom anderen einzusehen. An dem genauen *Inhalt* dieser normativen Erwartung bestimmt sich somit deckungsgleich auch, was angemessene Weisen darstellen, epistemische Gründe für das Vertrauen zu haben.

Diese Idee können wir anhand des oben genannten Fall defizienten Vertrauens testen. Im Falle des Ladenbesitzers, der seine Mitarbeitenden ausspioniert, scheint unmittelbar klar, dass diese wohl kaum ihrem Chef gegenüber erwarten, dass er sie aufgrund *solcherart* gewonnener epistemische Gründe als vertrauenswürdig ansieht. Vielmehr ist davon auszugehen, dass sie im Normalfall erwarten, dass er ihnen allein deshalb vertraut, da es keine bedeutenden Gründe gibt, misstrauisch ihnen gegenüber zu sein. Auch hier gilt natürlich wieder, dass, wie genau sich der normative Anspruch des Vertrauenspartners gestaltet, die Einsicht welcher Gründe genau er erwartet, von Fall zu Fall unterschiedlich sein kann. An dieser Stelle geht es uns jedoch darum, prinzipiell die Idee zu formulieren, dass sich die Grenze zwischen defizienten Vertrauensformen und unproblematischen Vertrauensformen genau an dem Inhalt des normativen Anspruches des Vertrauenspartners ergibt.

Um noch einen Schritt weiter zu gehen, ist zudem darauf aufmerksam zu machen, dass sich, so scheint mir, die normative Erwartung des Vertrauenspartners nicht nur auf die Einsicht in bestimmte Gründe erstreckt, sondern zuweilen auch auf den spezifischen *Modus*, in dem sich der Vertrauende auf diese Gründe zu beziehen hat. An dieser Stelle, meine ich, lohnt es sich deshalb einen Bogen zum vorangegangenen 8. Kapitel und dem Thema des habitualisierten Vertrauens zu schlagen. Es scheint mir nämlich, dass, insbesondere in intimen Vertrauensbeziehungen wie Freundschaften, die Anerkennung der Vertrauenswürdigkeit des anderen auch in einem bestimmten *Modus*, nämlich gerade *in habitualisierter Form* erwartet wird, der sich im Laufe der gemeinsamen Interaktionserfahrung auszubilden hat. Einem guten Freund etwa, mit dem es eine gemeinsame Geschichte wiederholter positiver Erfahrungen wechselseitigen Vertrauens und erprobter Vertrauenswürdigkeit gibt, kann im Normalfall nicht nur ohne Zögern vertraut werden. Mehr noch, in solchen intimen Beziehungsformen ist eine solche *habitualisierte* Vertrauenseinstellung gar wechselseitig voneinander *normativ verlangt*. Man erwartet vom Freund gewissermaßen, dass er das Vertrauen in einen ausreichend *verinnerlicht* hat. Die Erwartung an ihn besteht also darin, dass er die Gründe für das Vertrauen in den anderen, die sich aus den vergangenen Interaktionserfahrungen ergeben, bereits indirekt in einer *robusten* Haltung verdichtet hat. So erklärt sich auch, warum es bereits Empörung nachsichziehen oder auf andere Weise eine Störung der Beziehung zum anderen bedeuten kann, wenn jemand nur *zögert* bevor er seinem Freund vertraut. Auch wenn ein Freund sich letztendlich doch dazu durchringen mag, dem anderen zu vertrauen, bedeutet sein Zaudern bereits eine normative Devianz gegenüber den eigentlichen normativen Anforderungen qua freundschaftlicher Beziehung zum anderen.

Wenn ein Freund derart häsitiert, mögen wir uns mit einem Appell wie „Habe ich dich jemals enttäuscht?“ an ihn wenden. Man kann den Sinn dieses Appells so interpretieren, dass mit ihm unterstrichen werden soll, dass eine *berechtigte, da epistemisch abgesicherte,* normative Erwartung an den Freund besteht, als vertrauenswürdig angesehen zu werden. Der Aufruf an den Freund verweist dabei auf vergangene Erfahrungen, in denen man sich bereits als vertrauenswürdig erwiesen hat. Diese unterfüttern, dass es *gerechtfertigt* ist, dem Freund zu vertrauen, und dieser es damit auch zurecht erwarten kann.[407]

Oben sprachen wir in Bezug auf Vertrauen in seinem habitualisierten Modus vorrangig in Form seiner Praxis. In dem Kontext, wo das Vertrauen als anerkennende Haltung gegenüber einer bestimmten Person habitualisiert ist, bietet es sich an, Vertrauen vielmehr als einen persönlichen *Habitus* zu bezeichnen. Der Begriff des Habitus soll hier nur unterstreichen, dass man es mit einer *individuell eingewöhnten* Haltung zu tun hat, mit der man sich auf eine konkrete andere Person bezieht. Entscheidend ist dabei, dass man mit dieser anderen Person eine gemeinsame Geschichte hat, die in der habitualisierten Einstellung ihr gegenüber sedimentiert ist

Alles in allem haben wir damit Folgendes gezeigt: Wir sind von dem Gedanken ausgegangen, dass es sich bei Vertrauen um eine Anerkennung einer Person mit einer berechtigten normativen Erwartung gegenüber dem Vertrauenden handelt. Diese normative Erwartung des Vertrauensnehmers wurde präziser charakterisiert. Sie zielt nicht *einfach nur* darauf ab, dass man als vertrauenswürdig angesehen wird, sondern darauf, dass dies aus *Einsicht in bestimmte epistemische Gründe* geschieht, die für die eigene Vertrauenswürdigkeit sprechen. Damit zeigt sich ein innerer Zusammenhang zwischen Anerkennung und epistemischen Gründen. Weiterhin bestimmt sich auf dieser Idee aufbauend, ausgehend von dem genauen Inhalt dieser normativen Erwartung, auch, was angemessene Weisen sind, Vertrauen auf epistemische Gründe zu stützen. Der Vertrauenspartner erwartet etwa schlichtweg nicht, dass der Vertrauende ihm auf Grundlage solcher Gründe vertraut, die er durch Ausspionieren gewonnen hat. Fälle defizienten Vertrauens sind mit dem hier vorgestellten Ansatz also klar ausgeschlossen. Zuletzt haben wir darauf aufmerksam gemacht, dass insbesondere in intimen, auf einer gemeinsamen Geschichte beruhenden, Beziehungen, vom anderen auch ein bestimmter Modus erwartet wird, in dem er sich auf die epistemischen Gründe für sein Vertrauen bezieht. Demnach erwartet man etwa von einem

407Es ist in solchen intimen Beziehungen wie Freundschaften deshalb auch unangemessen, mit statistischen Gründen auf die Vertrauenswürdigkeit des anderen zu schließen. Ich vertraue dem Freund nicht, da ich es unter den gegebenen Bedingungen nach allgemeinen statistischen Zusammenhängen als wahrscheinlich ansehe, dass eine Person wie er sich vertrauenswürdig verhält. Ich vertraue ihm, da er im konkreten sich mir gegenüber wiederholt als vertrauenswürdig erwiesen hat.

Freund, dass dieser auf Grundlage der wiederholten positiven gemeinsamen Erfahrungen einen stabilen Habitus des Vertrauens ausbildet, in dem die epistemischen Gründe für sein Vertrauen gewissermaßen inkorporiert sind.

Mit diesem Vorschlag ließe sich im Übrigen auch ein weiterer Problemaspekt angehen, der zuweilen als Hindernis zur Versöhnung der interpersonellen und epistemischen Perspektive auf Vertrauen genannt wird. Wir können dies das Problem der Gründe falscher Art nennen. Dieses spielt bei Lackeys Formulierung des Dilemmas eine Rolle, das wir in der Einleitung dieses Kapitels geschildert haben, wird darüber hinaus aber auch ausdrücklich etwa von Thomas Simpson angesprochen.[408] Das grundlegende Problem lässt sich folgendermaßen skizzieren: Oben sprachen wir im Zusammenhang mit dem Kognitivismus des Vertrauens davon, dass Vertrauen, insofern es eine Überzeugung beinhaltet, nur durch epistemische Gründe als Gründe richtiger Art gestützt werden kann. Epistemische Gründe sind die richtige Kategorie an Gründen für Vertrauen. In interpersoneller Hinsicht aber verhält man sich im Vertrauen wesentlich angesichts einer *normativen* Anforderung, die von der anderen Person ausgeht. Die Frage stellt sich dann aber, wie man durch Akzeptieren einer normativen Anforderung zu einer Überzeugung gelangen kann, wenn Überzeugung grundlegend epistemisch begründet sein müssen. Immerhin, so die Überlegung, kann man doch nicht nur deshalb eine Überzeugung haben, da dies eben normativ gefordert sei, ebensowenig, wie man jemandem sinnvollerweise *befehlen* kann, eine bestimmte Überzeugung zu haben, etwa die, dass es in Köln freilebende Einhörner gibt.

Dieses Paradox scheint sich jedoch mit Hilfe unseres Vorschlages auflösen zu lassen. Die Lösung besteht darin zu erkennen, dass sich die normative Erwartung des Vertrauenspartners gerade *immer schon* auf epistemische Gründe bezieht. Die normative Erwartung des Vertrauenspartners besteht ja gerade darin, die *epistemischen Gründe einzusehen*, die für seine Vertrauenswürdigkeit sprechen. Damit besteht zumindest *kein prinzipieller* Gegensatz zwischen den praktischen und theoretischen Anforderungen in Bezug auf Vertrauen.

Zuletzt sei noch eine Bemerkung zur zuweilen geäußerten Kritik an der bekannten *evidential view* angebracht. Diese ist vor dem Hintergrund des hier vertretenen Ansatzes mit Bedacht einzuordnen. Wie wir oben in der Diskussion um Richard Moran gesehen haben, gibt es eine gerechtfertigte Kritik an der *evidential view*, da in dieser Vertrauen nicht als eine Haltung personaler Anerkennung verstanden werden kann. Daraus nun darf jedoch lediglich nicht der radikale Schluss gezogen werden, dass Vertrauen gar keinen Bezug mehr zu epistemischen

408Simpson 2018, 6; Vgl. Zagzebski 2012, 99-102.

Gründen aufweist. Damit wäre eine falsche und problematische Disjunktion zwischen der interpersonellen und epistemischen Seite von Vertrauen aufgemacht, in die sich der Vertrauensdiskurs zum Teil verstrickt hat.[409] Die Kritik an der *evidential view* muss deshalb vielmehr als eine Kritik an einer epistemisch *reduzierten* Betrachtung von Vertrauen verstanden werden. Vertrauen hat zwar auch eine epistemische Seite, ist jedoch zugleich mehr als dies.

Oben haben wir in der rein epistemischen Betrachtung von Vertrauen mit Rückgriff auf eine Idee von Pamela Hieronymi davon gesprochen, dass Vertrauen als Antwort auf die Frage rekonstruiert werden kann, ob die andere Person vertrauenswürdig ist. Dies stimmt zwar: Vertrauen stellt *auch* eine Antwort auf diese Frage dar und impliziert in dieser Hinsicht auch eine kognitive Überzeugung. Allerdings kann Vertrauen nicht darauf reduziert werden. Es ist zugleich mehr als das. Dies ist das eigentliche Problem der *evidential view,* die Vertrauen *lediglich* als Beantwortung der Frage verständlich macht, *ob es der Fall ist, dass* die andere Person vertrauenswürdig ist. Daraus ergibt sich, dass die Überlegung zur Beantwortung dieser Frage *allein* darauf abzielt eine möglichst korrekte Überzeugung über die andere Person zu bilden. Dies wäre eine rein theoretische Überlegung, deren Ergebnis eben kein Vertrauen in die andere Person darstellen kann.

Kann man dennoch Hieronymis Vorschlag, dass sich verschiedene mentale Zustände anhand einer charakteristischen Frage individuieren auf Vertrauen anwenden, ohne dieses zu einem epistemischen Phänomen zu reduzieren? Mein Vorschlag diesbezüglich lautet, dass sich Vertrauen in eine Person schlichtweg als Antwort auf die Frage konstituiert „Ist die andere Person als vertrauenswürdig *anzuerkennen*?“ An dieser Stelle ist erneut darauf hinzuweisen, dass dies nicht bedeutet, dass man nur insofern vertraut, als man such *ausdrücklich* mit dieser Frage auseinandergesetzt hat, um zu einer positiven Antwort zu kommen. Gemeint ist vielmehr, dass sich Vertrauen als eine Antwort auf eine solche Frage rekonstruieren ließe. Der Vorzug dieser Formulierung besteht nun darin, dass sie ausdrücklich berücksichtigt, dass interpersonelles Vertrauen letztlich als anerkennende Haltung verstanden werden muss. Aber wie wir gerade gesehen haben, wird diese Frage gerade unter Bezug auf epistemische Gründe beantwortet, die dafür sprechen, dass die andere Person vertrauenswürdig ist. Vertrauen bezieht sich damit, ebenso wie die bloße Überzeugung in die Vertrauenswürdigkeit der anderen Person, auf epistemische Gründe, bestimmte Evidenzen, die für die Vertrauenswürdigkeit der anderen Person, oder zumindest nicht gegen diese sprechen. Es bezieht sich auf diese jedoch im spezifischen *Lichte* einer nicht bloß kognitiven Frage nach der Wahrheit einer Überzeugung, sondern im Lichte der Frage, ob

409Siehe hierzu ausführlich: Kaminski 2017; 2020.

die andere Person als vertrauenswürdig *anzuerkennen* ist.

9.2 Epistemische Permissivität

Bisher haben wir nachgewiesen, dass sich Vertrauen, auch und gerade, da wir es als Anerkennung verstehen, an epistemischen Gründen orientiert. Zu beachten ist nun jedoch, dass wir in der gesamten obigen Diskussion im dritten und vierten Teil zwar wiederholt herausgearbeitet haben, dass sich Vertrauen auf epistemische Gründe und damit auf Evidenzen stützt, die für die Vertrauenswürdigkeit einer anderen Person sprechen. Damit haben wir aber womöglich suggeriert, dass Evidenzen ausschließlich einen gegenüber dem Vertrauen vorgängigen Charakter haben, und nicht umgekehrt auch Vertrauen einen Einfluss darauf hat, wie wir Evidenzen bewerten. Gegenüber dieser Sichtweise sind jedoch Zweifel anzumelden.

In der Literatur wurde an verschiedenen Stellen darauf hingewiesen, dass auch bei gleichem Kenntnisstand der epistemischen Faktenlage Akteure in ihrem Vertrauen- bzw. Misstrauensverhalten voneinander abweichen können. Und hierbei scheint insbesondere der Umstand, dass man sich in einer intimen Vertrauensbeziehung mit einer Person befindet, einen tiefgreifenden Unterschied machen zu können. Variieren wir zur Veranschaulichung wiederum das uns bekannte Beispiel von Judith Baker: Die Staatsanwaltschaft beschuldigt hierbei auf Grund der Beweislage eine Freundin von mir einer Straftat. Ich hingegen glaube der Versicherung meiner Freundin, dass sie unschuldig sei, da ich ihr vertraue. Natürlich wäre in diesem Fall mein Urteil nicht vollkommen epistemisch unbegründet, sondern stützt sich auf Erfahrungen, die ich mit meiner Freundin gemacht habe: Und zwar der Erfahrung, dass sie sich im Allgemeinen als vertrauenswürdig erwiesen hat. Stellen wir uns nun aber vor, der Staatsanwaltschaft sei meine Freundin schon lange verdächtig gewesen. Sie habe sie deswegen schon länger ausspionieren lassen und ihr Verhalten im Privatleben unter Beobachtung gestellt. Dabei wurden auch ihre Interaktionen mit mir dokumentiert und durchaus festgestellt, dass sie im Allgemeinen vertrauenswürdig eingestellt ist. Setzen wir also voraus, dass die Staatsanwaltschaft und ich im Grunde die gleichen relevanten Informationen über meine Freundin haben. Auf Grundlage einer nüchternen Betrachtung der Faktenlage jedoch, mag der Staatsanwalt insgesamt zu der Einschätzung kommen, dass meine Freundin schuldig ist, wohingegen ich an ihrer Unschuld festhalte.

Man beachte, dass der Unterschied zwischen den beiden epistemischen Einschätzungen letztendlich nicht nur an einer isolierten Einzelüberzeugung festzumachen ist. Die Position des Staatsanwaltes unterscheidet sich nicht bloß darin von meiner, dass er von der Schuld meiner Freundin überzeugt ist, während ich an ihrer Unschuld festhalte. Vielmehr geht dieser Unterschied in der Konsequenz damit einher, dass wir uns ein jeweils *strukturell* unterschiedliches Bild von der epistemischen *Gesamtsituation* machen. Das Vertrauen in meine Freundin hat die Implikation, dass ich die Fakten, die der Staatsanwalt als Beweis für die Schuld der Freundin ansieht, für mich bestenfalls den Status von *vermeintlichen* Beweisen haben können. Der Unterschied zwischen Staatsanwalt und Freund besteht damit in strukturell verschiedenen Deutungen der identischen Faktenlage. Vertrauen zeigt sich damit nicht einfach nur als Resultante bestimmter Evidenzen, sondern hat selber einen *strukturbildenden Effekt* auf die Bewertung der epistemischen Gesamtsituation. Dies soll nicht heißen, dass mit Vertrauen bestimmte Tatsachen einfach ignoriert werden. Wie gesagt, die rohe Faktenlage ist auch in dem Beispiel sowohl für den beobachtenden Staatsanwalt als auch für mich identisch. Die Ausgangslage wird lediglich anders *bewertet*, und in meinem Fall im Lichte meines Vertrauens zur Freundin betrachtet.[410]

Dieser Gedanke mag nun aber dazu verführen anzunehmen, dass Vertrauen letztendlich doch epistemisch willkürlich bleibt. Wenn die Idee einer rein objektiven Bewertung von Evidenzen abwegig ist, und Evidenzen vielmehr selber stets im Lichte bereits bestehender Vertrauens- oder Misstrauenseinstellungen vorgenommen werden, drängt sich das problematische Bild auf, dass sich Vertrauen letztendlich zirkulär stets selber rechtfertigen kann, und damit letztendlich doch zu einer rein arbiträren, nicht rational entscheidbaren Angelegenheit wird. So beispielsweise spitzt Lagerspetz das Problem zu, das sich aus dem obigen Fall ergeben mag.[411] Wir werden hingegen darlegen, dass sich dieses Problem nicht wirklich derart scharf stellt. Demgegenüber scheint der Fall aber auf einen anderen interessanten Umstand aufmerksam zu machen, nämlich den, dass es unter Umständen offen sein kann, ob Vertrauen oder Misstrauen in einer gegebenen Situation die epistemisch rationale Haltung ist. Diese Offenheit ist aber nicht einfach nur ein Defizit, sondern scheint gar Voraussetzung eines notwendigen Moments im Vertrauen, nämlich des Umstandes, dass sich der Vertrauende angesichts dieser Offenheit zu *verhalten* hat.

Sortieren wir noch einmal in Ruhe, wie der obige Fall zu verstehen ist. Zunächst ist festzuhalten, dass zugestanden werden kann, dass das Vertrauen in meine Freundin aus meiner eigenen Sicht als vernünftig verstanden wird. Als Vertrauender verstehe ich schließlich, wie wir

410Vgl. Lahno 2002, 138.
411Vgl. Lagerspetz 1998, 86.

oben herausgearbeitet haben, mein Vertrauen nicht einfach als willkürlich oder irrational, sondern als epistemisch angemessen. Genau dies hatten wir oben festgehalten, als wir von einem inneren Anspruch auf Situationsangemessenheit sprachen. Und genau dies verlangt ja auch die Freundin von mir: Sie möchte, dass ich *einsehe*, dass sie unschuldig ist und nicht etwa, dass ich mich in eine Selbsttäuschung diesbezüglich begebe.

Dass das vorgängige Vertrauen nun selber die Bewertung der epistemischen Evidenzen aus Sicht des Vertrauenden bestimmt, läuft aber nicht notwendig darauf hinaus, dass objektiv *unentscheidbar* ist, wann Vertrauen rational und wann es irrational ist. Es gibt weiterhin eindeutige Fälle, in denen dies klar ist: Stellen wir uns etwa vor, die Beweise gegen meine Freundin seien vielfach und unabweisbar. In diesem Fall ist klar, dass es ab einem gewissen Punkt eindeutig irrational wäre, weiterhin an ihrer Unschuld festzuhalten. Keine Frage: Ein festes und unbeirrbares Vertrauen zu einer bestimmten Person kann im konkreten Fall durchaus auf Kosten der epistemischen Rationalität gehen. Dies mag der Vertrauende aus seiner Perspektive zwar vielleicht anders sehen, in diesem Fall täuscht er sich aber. Dass Personen in ihrer Bewertung der epistemischen Situation voneinander abweichen, muss trivialerweise nicht bedeuten, dass gar nicht entscheidbar ist, ob es in der gegebenen Situation rational oder irrational ist zu vertrauen. Von daher wäre es ein *non sequitur* aus dem obigen Fall zu schließen, dass Vertrauen letztendlich doch eine epistemisch arbiträre Angelegenheit wäre, nur, da hier zwei Personen in ihrer epistemischen Einschätzung voneinander abweichen.

Was der obige Fall aber vielmehr aufzeigt, ist, dass es auch solche Situationen geben mag, in denen nicht klar entscheidbar ist, ob Vertrauen oder Misstrauen rational ist. Er weist darauf hin, dass die rohen Fakten als solche nicht eindeutig determinieren, was zu glauben rational, und was irrational wäre. Was in einer gegebenen Situation rational zu glauben sei, mag also durch dieselbe unterdeterminiert sein. Deshalb gibt es immer auch einen gewissen Spielraum für rationale Überzeugung, mit einem Begriff von Katherine Hawley auch als epistemische Permissivität zu bezeichnen.[412] Dieser erlaubt, dass man zugestehen kann, dass sich die beiden Perspektiven in dem Beispiel, zwischen der Staatsanwaltschaft und mir, im Rahmen eines *vernünftigen Dissenses* bewegen. Man kann zugestehen, dass der Staatsanwalt und der Freund in dem obigen Beispiel im Rahmen dieses rationalen Spielraumes eine unterschiedliche Gesamtbewertung der epistemischen Situation vornehmen. Dabei kann man sich ebenso vorstellen, dass sich beiden Seiten des Umstandes bewusst sind, dass sie einen vernünftigen Dissens zueinander haben. So könnte ich die

412Hawley 2014b, 2040.

Position der Staatsanwaltschaft nicht einfach als falsch abtun, sondern womöglich nachvollziehen ohne sie zu teilen. Umgekehrt mag auch die Staatsanwaltschaft dies über meine Ansicht denken.

In jedem Fall gilt, dass auch wenn mich die Vertrauensbeziehung zu meiner Freundin, offenbar in meiner Ansicht beeinflusst, sodass ich von der alternativen, epistemisch vollkommen nachvollziehbaren Position des Staatsanwaltes abweiche, ich mich deswegen nicht einfach irrational verhalte. Der Einfluss der Beziehung ist nicht einfach *nur* ein Vorurteil, das zwingend auf Kosten der epistemischen Rationalität geht. Mein Vertrauen ist insofern nicht einfach im schroffen Widerspruch zur Wahrheitsfindung. Vielmehr begründet das Verhältnis zu meiner Freundin einen bestimmten rationalen Zugriff auf die gegebene Faktenlage, der durch dessen Unterdeterminiertheit zugelassen wird. Die unterschiedlichen Schlussfolgerungen zu denen ich und die Staatsanwaltschaft kommen, mögen sich deshalb daraus erklären, dass uns unterschiedliche Sachen in diesem Fall wichtig sind.[413] Solange wir hier aber zugestehen, dass man sich innerhalb eines vernünftigen Dissens bewegt, können beide Perspektiven als prinzipiell rational angesehen werden.

Neben dieser Permissivität als Voraussetzung für rationalen Dissens, wollen wir aber noch einen Schritt weiter gehen und einen letzten Gedanken andeuten. Die epistemische Permissivität, die prinzipielle Unterbestimmtheit der epistemischen Faktenlage, ist demnach vielleicht nicht nur ein akzidentielles Element in Vertrauenskontexten, sondern vielleicht sogar eine notwendige Voraussetzung, um den *persönlich* anerkennenden Aspekt von Vertrauen verständlich machen zu können.

Die epistemische Permissivität bedeutet schließlich, dass sich der Vertrauender angesichts der epistemischen Faktenlage irreduzibel *verhalten*, Stellung beziehen muss. Noch einmal: Dies ist dabei nicht im radikal dezisionistischen, willkürlichen Sinne zu verstehen. Schließlich verhält sich der Vertrauende stets im Lichte epistemischer Gründe und mit dem Anspruch epistemische vernünftig zu sein. Und dennoch: Aufgrund der Unterdeterminiertheit der epistemischen Situation verbleibt im Vertrauen stets ein gewisses voluntatives Moment. Es steht einerseits zwar unter dem Anspruch epistemischer Rationalität. Andererseits muss der Vertrauende angesichts der Undeterminiertheit der epistemischen Situation irreduzibel Stellung beziehen. Dieses Moment des Verhaltens scheint nun aber gerade notwendig, um zu erklären, wie Vertrauen bzw. Misstrauen auch als Form der persönlichen Gestaltung der Beziehung zu einer anderen Person verstanden werden kann. Indem sich der Vertrauende angesichts der Unbestimmtheit der epistemischen Situation verhält, verhält er sich damit nämlich immer auch gegenüber der konkreten anderen

413 Siehe Lagerspetz' Bemerkung hierzu: „The court wants to close a case; I am trying to see what I am to make of our friendship in the light of the charges brought against my friend" (Lagerspetz 1998, 91).

Person.

10. Schluss: Vertrauen an der Schnittstelle zwischen Ethik und Epistemologie

Interpersonelles Vertrauen, das ist eine der Hauptlehren aus dieser Arbeit, hat zwar auch eine epistemische Seite, ist zugleich aber nicht auf diese zu reduzieren. Vertrauen in eine Person stellt nicht einfach nur eine bestimmte kognitive Einsicht dar. Man verfehlt entsprechend auch die volle Bedeutung einer Frage wie „Soll ich meinem Freund/ der Regierung/ den Impfherstellern vertrauen?“, wenn man sie als bloß theoretische Frage über die Faktenlage auffasst. Diese Fragen sind nicht bloß als Fragen nach dem Schema „Ist es der Fall, dass p“ aufzufassen. Damit würden sie nicht nur von der normativen Situierung von Vertrauen abstrahieren, sondern auch dessen anerkennungstheoretische Seite ausblenden. Es spielt dabei, wie wir gesehen haben, keine Rolle, im Lichte welchen Zweckes genau ein Akteur an der Frage interessiert ist, ob die andere Person vertrauenswürdig ist: Sei es, dass ihm am langfristigen Nutzen einer funktionierenden Kooperationsbeziehung gelegen ist, sei es, dass er daran interessiert ist zu erfahren, ob wahr ist, was sie sagt, etc. Insofern er allein eine kognitive Überzeugung hat, dass die andere Person vertrauenswürdig ist und aus diesem Grund bereit ist ein Risiko einzugehen, vertraut er noch nicht *ihr*. Die Bereitschaft zu Vertrauen ist schließlich irreduzibel als Bereitschaft zu verstehen, eine Person als vertrauenswürdig *anzuerkennen*.

Erst mit der Anerkennungstheorie haben wir deshalb den kategorialen Zugriffsrahmen erarbeitet, der die volle Signifikanz und Tragweite des Vertrauensphänomens angemessen in den Blick nimmt. Sie erlaubt es die Bedeutung, die Vertrauen und Misstrauen im Kontext und zur Gestaltung interpersoneller Beziehungen hat zu erfassen, und emotionale Reaktionen auf Vertrauen und Misstrauen zu erklären. Damit zeigt sich deutlich, dass Vertrauen und Misstrauen mehr sind als Ausdruck eines Kenntnisstandes über das Faktum, ob die andere Person vertrauenswürdig ist oder nicht. Am Vertrauen entscheidet sich vielmehr, ob die andere Person ernst genommen, ob ihr ein bestimmter normativer Status zugeschrieben wird. Mit diesem Verständnis von Vertrauen ist somit der Horizont dafür geöffnet, die genuin interpersonelle Bedeutung von Vertrauen und Misstrauen herauszustellen.

Wir haben darüber hinaus gezeigt, dass es einen inhärenten Zusammenhang zwischen

Vertrauen als Anerkennung und dem Anspruch darauf gibt, dass dieses epistemisch begründet ist. Die interpersonelle und die epistemische Betrachtung von Vertrauen stehen deshalb nicht bloß dichotom nebeneinander. Vielmehr weist eine genauere Analyse des Anerkennungsbegriffes von sich aus bereits auf epistemische Bezugspunkte. Einer Person zu vertrauen bedeutet demnach gerade, ihrem normativen Anspruch darauf nachzukommen *einzusehen*, dass sie vertrauenswürdig ist. Nur insofern kommt schließlich dem Charakter des Ernstnehmens vom Vertrauen auch die entscheidende Ernst*haftigkeit* zu.

Die Frage, wer ernst genommen und wer nicht ernst genommen, wem also vertraut und wem nicht vertraut wird, ist dabei als eine zentrale ethische Frage aufzufassen. Am Vertrauen und Misstrauen entscheidet sich nicht nur, wer Anerkennung erfährt, sondern auch ganz praktisch, wem dadurch bestimmte Handlungsoptionen gewährt, und wem verschlossen werden. Gerade aus diesem Grunde ist es auch von enormer ethischer Bedeutung, dass Vertrauen und Misstrauen epistemisch vernünftig vollzogen werden. Hinzu kommt, dass Fragen des Vertrauens und Misstrauens erkennbare Konsequenzen für Dritte oder gar das soziale Zusammenleben als Ganzes haben. Problematische Formen des Vertrauens in Missinformationen und deren schädliche Auswirkungen auf den öffentlichen Diskurs und demokratische Entscheidungsprozesse seien hier nur als ein aktuelles Beispiel genannt.

Auch wenn sich natürlich immer am Einzelfall entscheiden muss, was genau als unangemessenes, zu undifferenziertes oder zu dogmatisches Vertrauen bzw. Misstrauen zählt, gilt zumindest formal formuliert die Maxime, dass Vertrauen stets angemessen reflektiert und differenziert werden sollte. Gerade aufgrund der weitreichenden anerkennungsbezogenen und pragmatischen Implikation von Vertrauen und Misstrauen wird deutlich, dass deren *epistemisch* rationale Ausübung bzw. das Schaffen von Bedingungen, unter denen sie epistemisch vernünftig kontrolliert ausgeübt werden können, eine *ethisch* bedeutsame Aufgabe darstellt. Vertrauen stellt somit insgesamt ein Phänomen dar, das weder einfach nur in ethisch-praktischer oder epistemisch-theoretischer Hinsicht zu betrachten ist, sondern das gerade darauf hinweist, wie diese beiden Seiten auch notwendigerweise miteinander verknüpft sind.

Bibliographie

Adler, Jonathan. „Epistemological problems of testimony“. In *The Stanford Encyclopedia of Philosophy (Winter 2006 Edition)*, herausgegeben von Edward N. Zalta, 2006. URL = <https://plato.stanford.edu/archives/win2006/entries/testimony-episprob/>, abgerufen am 28.09.2020.

Alkemeyer, Thomas/ Schürmann,Volker/ Volbers, Jörg, Hrsg. *Praxis denken*. Wiesbaden: VS, 2015.

Annas, Julia. „Virtue ethics“. In *The Oxford Handbook of Ethical Theory*, herausgegeben von David Copp, 515–36. Oxford: Oxford University Press, 2006.

Anscombe, Gertrude Elizabeth Margaret. *Intention*. 2. Aufl. Harvard: Harvard University Press, 2000.

———.„What is it to believe someone?“ In *Faith in a Hard Ground: Essays on Religion, Philosophyand Ethics*, herausgegeben von Mary Geach und Luke Gormally, 1–10. Exeter: Imprint Academics, 2008.

Aristoteles. *Nikomachische Ethik*. Stuttgart: Reclam, 1969.

———. *Rhetorik*. München: Wilhelm Fink, 1980.

Austin, John L. „Other minds“. In *Knowledge: Readings in Contemporary Epistemology*, herausgegeben von Sven Bernecker und Fred I. Dretske, 339-346. Oxford: Oxford University Press, 2000.

Baier, Annette. „Trust and antitrust“. *ethics* 96, Nr. 2 (1986): 231–60.

Baker, Judith. „Trust and rationality“. *Pacific philosophical quarterly* 68, Nr. 1 (1987): 1–13.

Baliga, Sandeep, und Tomas Sjöström. „The hobbesian trap“. In *The Oxford Handbook of the Economics of Peace and Conflict*, herausgegeben von Michelle R. Garfinkel und Stergios

Skaperdas, 91–109. Oxford: Oxford UP, 2010.

Berg, Joyce/ Dickhaut, John/ McCabe, Kevin. „Trust, reciprocity, and social history." *Games and Economic Behavior* 10, Nr.1 (1995): 122–142.

Brugger, Philipp/ Hasenclever, Andreas/ Kasten, Lukas. „Vertrauen lohnt sich: Über Gegenstand und Potential eines vernachlässigten Konzepts in den Internationalen Beziehungen". *Zeitschrift für internationale Beziehungen* 20, Nr. 2 (2013): 65–104.

Budnik, Christian. „Gründe für Vertrauen, Vertrauenswürdigkeit und Kompetenz". *Deutsche Zeitschrift für Philosophie* 64, Nr. 1 (2016): 103–18.

Burge, Tyler. „Content preservation". *Philosophical Issues* 6 (1995): 271–300.

Camerer, Colin/ Weigelt, Keith. „Experimental tests of a sequential reputation model". *Econometrica: Journal of the Econometric Society* 56, Nr. 1 (1988): 1–36.

Cavell, Stanley. *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy.* Oxford: Oxford University Press, 1979.

———.„Aesthetic problems of modern philosophy". In ders., *Must We Mean What We Say*, 68–90. Cambridge: Cambridge University Press, 2015.

Cetina, Karin Knorr/ Schatzki, Theodore R./ Von Savigny, Eike, Hrsg. *The Practice Turn in Contemporary Theory*. London: Routledge, 2001.

Chwaszcza, Christine. *Praktische Vernunft als vernünftige Praxis.* Weilerswist: Velbrück Wissenschaft, 2003.

Coady, Cecil Anthony John. *Testimony: A philosophical study*. Oxford: Clarendon Press, 1992.

Cogley, Zac. „Trust and the trickster problem". *Analytic Philosophy* 53, Nr. 1 (2012): 30–47.

Coleman, James S. *Foundations of Social Theory*. Harvard: Harvard University Press, 1994.

Cook, Karen S./ Hardin, Russell/ Levi, Margaret. *Cooperation Without Trust?* New York: Russell

Sage Foundation, 2005.

Cook, Karen S./ Santana, Jessica J. „Trust and Rational Choice“ In *The Oxford Handbook of Social and Political Trust* herausgegeben von Eric M. Uslaner, 253-78. Oxford: Oxford University Press, 2018.

Darwall, Stephen. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Harvard: Harvard University Press, 2006.

———.„Moral obligation and accountability”. In *Oxford Studies in Metaethics*, 2.Bd., herausgegeben von Russ Shafer-Landau, 111-132. Oxford: Oxford University Press, 2007.

———.„Demystifying promises“. In *Promises and agreements: Philosophical essays*, herausgegeben von Hanoch Sheinman, 255–76. Oxford: Oxford University Press, 2011.

———.„Trust as a second-personal attitude (from the heart)“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 35–50. Oxford: Oxford University Press, 2017.

Dasgupta, Partha. „Trust as a commodity“. In *Trust: Making and Breaking Cooperative Relations*, herausgegeben von Diego Gambetta, 49:49–72. Cambridge, Mass: Basil Blackwell, 1988.

Davidson, Donald. „Actions, reasons, and causes“. *The journal of philosophy* 60, Nr. 23 (1963): 685–700.

Deutsch, Morton. „Cooperation and trust: Some theoretical notes“. In *Nebraska Symposium on Motivation*, herausgegeben von Marshall R. Jones, 275–320. Lincoln: Univer. Nebraska Press., 1962.

———. „Trust and suspicion“. *Journal of Conflict Resolution* 2, Nr. 4 (1958): 265–79.

Dewey, John. *Human Nature and Conduct: An Introdcution too Social Psychology.* New York: Cosimo, 2007.

Domenicucci, Jacopo/ Holton, Richard. „Trust as a two-place relation“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas W. Simpson, 149–60. Oxford: Oxford University Press, 2017.

Douskos, Christos. „Habit and intention“. *Philosophia 45*, Nr.3 (2017): 1129–1148 .

Dormandy, Katherine. „Exploitative epistemic trust“. In *Trust in Epistemology*, herausgegeben von Katherine Dormandy, 241–64. New York/Oxon: Taylor & Francis, 2020.

Dunning, David/ Anderson, Joanna E./ Schlösser, Thomas/ Ehlebracht, Daniel/ Fetchenhauer, Detlef . „Trust at zero acquaintance: More a matter of respect than expectation of reward.“ *Journal of Personality and Social Psychology* 107, Nr. 1 (2014): 122.

Endreß, Martin. *Vertrauen*. Bielefeld: transcript Verlag, 2002.

———. „Zur Struktur von ‚Grundvertrauen‘ und der Vertrauenssignatur in Gegenwartsgesellschaften“. In *Grundvertrauen: Hermeneutik eines Grenzphänomens*, herausgegeben von Ingolf U. Dalferth und Simon Peng-Keller. Leipzig: Evangelische Verlagsanstalt, 2013.

Enoch, David. „Authority and reason-giving“. *Philosophy and Phenomenological Research* 89, Nr. 2 (2014): 296–332.

———. „Giving practical reasons“. *The Philosopher's Imprint* 11, Nr. 4 (2011).

Faulkner, Paul. „IX—A Virtue Theory of Testimony“. In *Proceedings of the Aristotelian Society*, 114:189–211. Oxford University Press Oxford, UK, 2014.

———. *Knowledge on Trust*. Oxford/New York: Oxford University Press, 2011.

———. „On telling and trusting“. *Mind* 116, Nr. 464 (2007): 875–902.

———. „Replies to critics“. Abstracta, Special Issue VI, (2012): 117 – 137.

———. „The practical rationality of trust“. *Synthese* 191, Nr. 9 (2014): 1975–89.

———. „The problem of Trust“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 109–28. Oxford/New York: Oxford University Press, 2017.

Foot, Philippa. „Utilitarianism and the virtues“. *Mind* 94, Nr. 374 (1985): 196–209.

Frevert, Ute. *Vertrauensfragen: Eine Obsession der Moderne*. München: C.H.Beck, 2013.

Fricker, Elizabeth. „Against gullibility“. In *Knowing from words,* herausgegeben von Bimal K. Matilals und Arindam Chakrabarti, 125–61. Dordrecht: Kluwer Academic Publishers, 1994.

———. „Telling and trusting: Reductionism and anti-reductionism in the epistemology of testimony“. Mind 104, Nr. 414 (1995): 393–411.

———. „Second-hand knowledge“. *Philosophy and Phenomenological Research* 73, Nr. 3 (2006):

592–618.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford/New York: Oxford University Press, 2007.

Frost-Arnold, Karen. „The cognitive attitude of rational trust". *Synthese* 191, Nr. 9 (2014): 1957–74.

Gambetta, Diego. „Can we Trust Trust". In *Trust: Making and Breaking Cooperative Relations*, herausgegeben von Diego Gambetta, 13:213–37. Cambridge, Mass: Basil Blackwell, 1988.
———. *Trust: Making and breaking cooperative relations*. B. Blackwell New York, NY, 1988.

Geach, Peter T. „Good and evil". *Analysis* 17, Nr. 2 (1956): 33–42.

Gelfert, Axel. „Hume on testimony revisited". *History of Philosophy & Logical Analysis* 13, Nr.1 (2010): 60-75.
———. *Testimony: A Critical Introduction.* London/New Delhi/New York/Sydney: Blomsbury, 2014.

Gertken, Jan/ Kiesewetter, Benjamin. „The right and the wrong kind of reasons". *Philosophy Compass* 12, Nr. 5 (2017): e12412.

Gilbert, Margaret. „Scanlon on promissory obligation: The problem of promisees' rights". *The Journal of Philosophy* 101, Nr. 2 (2004): 83–109.

Goldberg, Sanford/ Henderson, David. „Monitoring and anti-reductionism in the epistemology of testimony". *Philosophy and Phenomenological Research* 72, Nr. 3 (2006): 576-593.

Govier, Trudy. *Dilemmas of Trust*. Montreal & Kingston/London/Ithaka: McGill-Queen's Press-MQUP, 1998.

Grice, Paul. „Meaning" *The Philosophical Review* 66, Nr. 3 (1957): 377-388.

Haase, Matthias. „Drei Formen der Ersten Person Plural". *Deutsche Zeitschrift für Philosophie* 55, Nr. 2 (2007): 225–43.

———. „For oneself and toward another: The puzzle about recognition", in: *Philosophical Topics* 42, Nr. 1 (2014): 113-152.

Hahn, Susanne. *Rationalität*. Paderborn: mentis, 2013.

Halbig, Christoph. „Normative Gründe. Neue Beiträge zur Theorie praktischer Vernunft (II)". *Deutsche Zeitschrift für Philosophie* 51, Nr. 1 (2003): 133–49.

Hanfling, Oswald. *Philosophy and Ordinary Language: The Bent and Genius of Our Tongue*. London/New York: Routledge, 2000.

Hardin, Russell. *Trust and Trustworthiness*. New York: Russell Sage Foundation, 2002.
———. „Trustworthiness". *Ethics* 107, Nr. 1 (1996): 26–42.

Harding, Matthew. „Responding to trust". *Ratio Juris* 24, Nr. 1 (2011): 75–87.

Hartmann, Martin. *Die Praxis des Vertrauens*. Frankfurt a.M.: Suhrkamp Verlag, 2011.

Hartmann, Martin/ Offe, Claus, Hrsg. *Vertrauen: die Grundlage des sozialen Zusammenhalts*. Frankfurt a.M.: Campus Verlag, 2001.

Hawley, Katherine. „Trust, distrust and commitment". *Noûs* 48, Nr. 1 (2014a): 1–20.
———. „Partiality and prejudice in trusting". *Synthese 191*, Nr. 9 (2014b): 2029–45.
———. „Trustworthy groups and organizations". In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 230–50. Oxford: Oxford UP, 2017.
———. *How to Be Trustworthy*. Oxford: Oxford University Press, 2019.

Heap, Shaun Hargreaves/ Varoufakis, Yanis. *Game Theory: a Critical Text*. London: Routledge, 1995.

Helm, Bennett W. „Responsibilty and Dignity: Strawsonian Themes". In *Morality and the Emotions*, herausgegeben von Carla Bagnoli, 217–34. Oxford/New York: Oxford University Press, 2011.
———. „Trust as a reactive attitude". *Oxford studies in agency and responsibility* 2 (2014): 187–215.
———. „Friendship". In *The Stanford Encyclopedia of Philosophy (Fall 2017 Edition)*,

herausgegeben von Edward N. Zalta, 2017. URL = <https://plato.stanford.edu/archives/fall2017/entries/friendship/>, abgerufen am 20.6.2020.

Hertzberg, Lars. „On the attitude of trust“. *Inquiry* 31, Nr. 3 (1988): 307–22.

Heuer, Ulrike. „Promising - Part 2“. *Philosophy Compass* 7, Nr. 12 (2012): 842–51.

Hieronymi, Pamela. „The reasons of trust“. *Australasian Journal of Philosophy* 86, Nr. 2 (2008): 213–36.

———. „The use of reasons in thought (and the use of earmarks in arguments)“. *Ethics* 124, Nr. 1 (2013): 114–27.

Hinchman, Edward S. „Telling as inviting to trust“. *Philosophy and Phenomenological Research* 70, Nr. 3 (2005): 562–87.

———. „Can trust itself ground a reason to believe the trusted?“ *Abstracta* 6, Nr. 6 (2012): 47-83.

———. „Assurance and warrant“. *Philosophers' Imprint* 14 (2014): 1–58.

———. „On the risks of resting assured: An assurance theory of trust“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 51–69. Oxford: Oxford UP, 2017.

Hoffmann, Thomas. „Lebensform – Natur, Begriff und Norm“. In *Normativität des Lebens – Normativität der Vernunft?*, herausgegeben von Markus Rothhaar und Martin Hähnel, 45–64. Berlin: De Gruyter, 2015.

Hollis, Martin. *Trust within Reason*. Cambridge: Cambridge University Press, 1998.

Holton, Richard. „Deciding to trust, coming to believe“. *Australasian Journal of Philosophy* 72, Nr. 1 (1994): 63–76.

Honneth, Axel, und Beate Rössler. „Einleitung: Von Person zu Person: Zur Moralität persönlicher Beziehungen“. In *Von Person zu Person: Zur Moralität persönlicher Beziehungen*, herausgegeben von Axel Honneth und Beate Rössler, 9–25. Frankfurt a.M.: Suhrkamp, 2008.

Horsburgh, Howard John Neate. „The ethics of trust“. *The Philosophical Quarterly* 10, Nr. 41 (1960): 343–54.

Hume, David. *A Treatise of Human Nature: Volume 2*. Herausgegeben von David Fate Norton und Mary J. Norton. Oxford: Oxford UP, 2007a.

———. *An Enquiry Concerning Human Understanding.* Herausgegeben von Peter Millican. Oxford: Oxford UP, 2007b.

———. *An Enquiry Concerning the Principles of Morals*. Herausgegeben von Jerome B. Schneewind. Indianapolis: Hackett Publishing Company, 1983.

Ikäheimo, Heikki, und Arto Laitinen. „Analyzing recognition: Identification, acknowledgement and recognitive attitudes towards persons“. In *Recognition and Power*, herausgegeben von Bert van den Brink und David Owen, 33–56. Cambridge: Cambridge University Press, 2007.

Jaeggi, Rahel. *Kritik von Lebensformen*. Berlin: Suhrkamp Verlag, 2014.

John Rawls. *A Theory of Justice*. Harvard: Belknap Press - Harvard University Press, 1999.

John, Stephen. „Epistemic trust and the ethics of science communication: Against transparency, openness, sincerity and honesty“. *Social Epistemology* 32, Nr. 2 (2018): 75–87.

Johnson, Branden B. „Exploring dimensionality in the origins of hazard-related trust“. *Journal of Risk Research* 2, Nr. 4 (1999): 325-354.

Jones, Karen. „‘But I was counting on you!’“. In *The Philosophy of Trust*, herausgegeben von Thomas Simpson und Paul Faulkner, 90–108. Oxford: Oxford University Press, 2017.

———. „Trust as an affective attitude“. *Ethics* 107, Nr. 1 (1996): 4–25.

———. „The politics of intellectual self-trust“. *Social Epistemology* 26, Nr. 2 (2012a): 237-251.

———. „Trustworthiness“. *Ethics* 123, Nr. 1 (2012b): 61–85.

Kaminski, Andreas. „Die Form von Vertrauen und ihre verwickelte Praxis“. In *Vertrauen und Transparenz – für ein neues Europa*, herausgegeben von Alfred Hirsch, Petar Bojanić, und Željko Radinković, 163–83. Belgrad: Institut für Philosophie und Gesellschaftstheorie Belgrad, Universität Belgrad, 2013.

———. „Hat Vertrauen Gründe oder ist Vertrauen ein Grund? Eine dialektische Tugendtheorie von Vertrauen und Vertrauenswürdigkeit“. In *Praxis und ›zweite Natur‹: Begründungsfiguren*

normativer Wirklichkeit in der Diskussion, herausgegeben von Jens Kertscher und Jan Müller, 167–88. Paderborn: mentis, 2017.

———. „Die verwickelte Einfachheit von Vertrauen und seine spekulative Struktur". Habilitationsschrift, Universität Marburg, 2020.

Keren, Arnon. „On the alleged perversity of the evidential view of testimony". *Analysis* 72, Nr. 4 (2012): 700–707.

———. „Trust and belief: a preemptive reasons account". *Synthese* 191, Nr. 12 (2014): 2593–2615.

Kern, Andrea. *Quellen des Wissens: Zum Begriff vernünftiger Erkenntnisfähigkeiten*. Frankfurt a.M.: Suhrkamp, 2006.

Kietzmann, Christian. *Handeln aus Gründen als praktisches Schließen*. Bd. 93. Freiburg im Breisgau: Verlag Herder GmbH, 2019.

Kolodny, Niko/ Wallace, R. Jay. „Promises and practices revisited". *Philosophy & Public Affairs* 31, Nr. 2 (2003): 119–54.

Korsgaard, Christine M. „Skepticism about practical reason". *The Journal of Philosophy* 83, Nr. 1 (1986): 5–25.

Lackey, Jennifer. *Learning from Words: Testimony as a Source of Knowledge*. Oxford/New York: Oxford University Press on Demand, 2008.

Lagerspetz, Olli. *Trust: The Tacit Demand*. Berlin/Heidelberg: Springer Science & Business Media, 1998.

Lahno, Bernd. „Ist Vertrauen eine rationale Erwartung?" In *Rationalität, Realismus, Revision. Vorträge des 3. internationalen Kongresses der Gesellschaft für analytische Philosophie vom 15. bis zum 18. September 1997 in München,* herausgegeben von Julian Nida-Rümelin, 308-316. Berlin: De Gruyter, 2000.

———. *Der Begriff des Vertrauens*. Paderborn: Mentis, 2002."

———. „On the emotional character of trust". *Ethical theory and moral practice* 4, Nr. 2 (2001): 171–89.

Løgstrup, Knud Ejler. *The Ethical Demand*. Notre Dame: University of Notre Dame Pess, 1997.

Longworth, Guy. „Faith in others“. *Abstracta* 6, Nr. 3 (2012): 6–32.

Löschke, Jörg. „Friendship, value and interpretation“. *Theoria* 83, Nr. 4 (2017): 319–40.

Luhmann, Niklas. *Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität*. 4. Aufl. Stuttgart: Lucius & Lucius, 2014.

Marusic, Berislav. *Evidence and Agency: Norms of Belief for Promising and Resolving*. Oxford: Oxford UP, 2015.

Mayer, Roger C./ Davis James H. „The effect of the performance appraisal system on trust for management: A field quasi-experiment“. *Journal of Applied Psychology* 84, Nr.1 (1999): 123-136.

McDowell, John. „Virtue and reason“. *The monist* 62, Nr. 3 (1979): 331–50.

———. „Might There Be External Reasons?“ In *World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, herausgegeben von J. E. J. Altham und Ross Harrison, 68–85. Cambridge: Cambridge University Press, 1995.

———. „Knowledge by hearsay“. In ders., *Meaning, Knowledge and Reality,* 414-443. Cambridge/London: Harvard University Press, 1998.

McGeer, Victoria. „Developing trust“. *Philosophical Explorations 5,* Nr.1 (2002): 21–38.

———.„Trust, hope and empowerment“. *Australasian Journal of Philosophy* 86, Nr. 2 (2008): 237–54.

McLeod, Carolyn. *Self-trust and Reproductive Autonomy*. Cambridge: MIT Press, 2002.

Michael Thompson. „What is it to wrong someone ? A puzzle about justice“. In *Reason and Value,* herausgegeben von Philip Pettit, Samuel Scheffler, Michael Smith, und Jay Wallace, 333–84. Oxford: Oxford University Press, 2003.

Miller, Jessica. „Trust in strangers, trust in friends“. *Philosophy in the Contemporary World* 10, Nr. 1 (2003): 17–22.

Moore, George E. „Moore's paradox“. In *G. E. Moore: Selected Writings,* herausgegeben von Thomas Baldwin, 207–212. London: Routledge, 1993.

Moran, Richard. „Gettin told and being believed“. *Philosophers' Imprint* 5, Nr. 5 (2005): 1-29.

Moyal-Sharrock, Danièle. *Understanding Wittgenstein's On Certainty.* Hampshire/New York: Palgrave Macmillan, 2004.

———. „Wittgenstein's grammar: Through thick and thin“. In *WITTGENSTEINIAN (adj.) Looking at the World from the Viewpoint of Wittgenstein's Philosophy*, herausgegeben von Shyam Wuppuluri und Newton da Costa, 39–54. Berlin/Heidelberg: Springer, 2020.

Müller, Anselm. „VIII—Mental Teleology“. In *Proceedings of the Aristotelian Society*, 92: 161–84. Oxford University Press Oxford, UK, 1992.

Müller, Jan. „» Lebendige Normativität «?“ In *Lebensform und Praxisform*, herausgegeben von Jens Kertscher und Jan Müller, 263–87. Paderborn: mentis, 2015.

Müller, Wolfgang. „Risiko und Ungewißheit“. In *Handwörterbuch der Betriebswirtschaft*, herausgegeben von, Werner Kern, Richard Köhler, Hans-Ulrich Küppe, Klaus v. Wysocki und Waldemar Wittmann, 3:3813–24. Stuttgart: Schäfer-Poeschel Verlag, 1993.

Nickel, Philip J. „Being pragmatic about trust“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 195–213. Oxford: Oxford University Press, 2017.

———. „Trust and obligation-ascription“. *Ethical theory and moral practice* 10, Nr. 3 (2007): 309–19.

———. „Trust and testimony“. *Pacific Philosophical Quarterly* 93, Nr. 3 (2012): 301–16.

Nida-Rümelin, Julian. *Philosophie und Lebensform*. Frankfurt a.M.: Suhrkam, 2009.

Offe, Claus. „How can we trust our fellow citizens“. In *Democracy and Trust*, herausgegeben von Mark E. Warren, 42–87. Cambridge: Cambridge University Press, 1999.

O'Neil, Collin. „Betraying trust“. In *The Philosophy of Trust*, herausgegeben von Paul Faulkner und Thomas Simpson, 70–89. Oxford: Oxford University Press, 2017.

O'Neil, Collin. „Lying, trust, and gratitude". *Philosophy & Public Affairs* 40, Nr. 4 (2012): 301–33.

O'Neill, Onora. *A Question of Trust: The BBC Reith Lectures 2002*. New York: Cambridge University Press, 2002.

Owens, David. „A simple theory of promising". *The Philosophical Review* 115, Nr. 1 (2006): 51–77.

———. *Shaping the Normative Landscape*. Oxford: Oxford University Press, 2012.

Peirce, Charles Sanders. „The Fixation of Belief". In *The Essential Peirce I*, herausgegeben von Nathan Houser und Christian Kloesel, 109-123. Indianapolis: Indiana University Press, 1992.

Pettit, Philip. „The cunning of trust". *Philosophy & Public Affairs* 24, Nr. 3 (1995): 202–25.

Pihlström, Sami. „A new look at Wittgenstein and pragmatism". *European Journal of Pragmatism and American Philosophy* 4, Nr. 2 (2012): 9–26.

Potter, Nancy Nyquist. *How Can I Be Trusted?: a Virtue Theory of Trustworthiness*. Oxford: Rowman & Littlefield, 2002.

Price, Henry Habberley. *Belief.* London: Allen & Unwin, 1969.

Prijic-Samarzija, Snjezana. „Trust and contextualism". *Acta Analytica* 22, Nr.2 (2007): 125-138.

Putnam, Robert D. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press, 1992.

Rabinowicz, Wlodek, und Toni Rønnow-Rasmussen. „The strike of the demon: On fitting pro-attitudes and value". *Ethics* 114, Nr. 3 (2004): 391–423.

Raz, Joseph. *Engaging Reason: On the Theory of Value and Action*. Oxford: Oxford University Press, 1999.

Ricoeur, Paul. *The Course of Recognition*. Bd. 2. Havard: Havard University Press, 2004.

Ripperger, Tanja. *Ökonomik des Vertrauens: Analyse eines Organisationsprinzips*. Tübingen: Mohr Siebeck, 2003.

Rödl, Sebastian. *Selbstbewußtsein*. Frankfurt a.M.: Suhrkamp, 2011.

Rønnow-Rasmussen, Toni/ Zimmerman, Michael J. *Recent Work on Intrinsic Value*. Luxemburg: Springer Science & Business Media, 2006.

Scanlon, Thomas. *What We Owe to Each Other*. Cambridge, Mass: Harvard UP, 1999.

Schäfer, Hilmar. *Praxistheorie: ein soziologisches Forschungsprogramm*. Bielefeld: transcript Verlag, 2016.

Scheffler, Samuel. „Agent-centred restrictions, rationality, and the virtues". *Mind, New Series* 94, Nr. 375 (1985): 409–19.

———. „Valuing". In *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*, herausgegeben von R. Jay Wallace, Rahul Kumar, und Samuel Freeman, 23–42. New York: Oxford University Press, 2011.

Simmel, Georg. *Soziologie: Untersuchungen über die Formen der Vergesellschaftung*. Herausgegeben von Otthein Ramstedt. Bd. 11. 24 Bde. Georg Simmel Gesamtausgabe. Frankfurt a.M.: Suhrkamp, 1992.

Simpson, Thomas W. „Trust, belief, and the second-personal". *Australasian Journal of Philosophy* 96, Nr. 3 (2018): 447–59.

———. „What is trust?" *Pacific Philosophical Quarterly* 93, Nr. 4 (2012): 550–69.

Smith, Michael. „The Humean theory of motivation". *Mind* 96, Nr. 381 (1987): 36–61.

Smolkin, Doran. „Puzzles about trust". *The Southern journal of philosophy* 46, Nr. 3 (2008): 431–49.

Stern, Robert. *The Radical Demand in Løgstrup's Ethics*. Oxford: Oxford University Press, 2019.

Strawson, Peter. „Freedom and resentment". In ders., *Studies in the philosophy of thought and action: 1968 British Academy Lectures,*71–96. London: Oxford University Press, 1968.

Sullivan, John L./ Transue, John E. „The psychological underpinnings of democracy: A selective review of research on political tolerance, interpersonal trust, and social capital“. *Annual Review of Psychology* 50, Nr. 1 (1999): 625–50.

Szabados, Bèla. „Wittgenstein on 'Mistrusting One's Own Belief'“. *Canadian Journal of Philosophy* 11, Nr. 4 (1981): 603 – 612.

Sztompka, Piotr. „Trust, distrust and two paradoxes of democracy“. *European Journal of Social Theory* 1, Nr. 1 (1998): 19–32.

Thompson, Christopher. „Trust Without Reliance“. *Ethical Theory and Moral Practice* 20, Nr. 3 (2017): 643–55.

Thompson, Michael. „What is it to wrong someone?: A puzzle about justice*“.* In *Reason and Value: Themes from the Moral Philosophy of Joseph Raz,* herausgegeben von R. Jay Wallace, Philip Pettit, Samuel Scheffler und Michael Smith, 333–84. Oxford: Clarendon Press, 2004.

———. *Life and Action: Elementary Structures of Practice and Practical Thought.* Cambridge: Harvard University Press, 2008.

Townley, C. *A Defense of Ignorance: Its Value for Knowers and Roles in Feminist and Social Epistemologie.* Lanham, MD: Lexington Books, 2011.

Velleman, J. David. „The guise of the good“. *Nous* 26, Nr. 1 (1992): 3–26.

Vitale, D. „A trust network model for social rights“. *Oxford Journal of Legal Studies* 38, Nr.4 (2018): 1–27.

Walker, Margaret Urban. *Moral repair: Reconstructing moral Relations after Wrongdoing*. Cambridge: Cambridge University Press, 2006.

Wallace, James D. *Norms and Practices*. Ithaca: Cornell University Press, 2009.

Wallace, R. Jay. „Dispassionate opprobrium: On blame and the reactive sentiments“. In *Reasons and recognition: Essays on the philosophy of TM Scanlon*, herausgegeben von R. Jay Wallace, Rahul Kumar, und Samuel Freeman, 348–72. Oxford: Oxford University Press,

2011.
———. *Responsibility and the Moral Sentiments*. Harvard: Harvard University Press, 1994.
———. „The deontic structure of morality“. In *Thinking about reasons. Essays in honour of jonathan dancy*, herausgegeben von David Bakhurst, Brad Hooker, und Margaret Olivia Little, 137–67. Oxford: Oxford University Press, 2013.
———. *The Moral Nexus*. Princeton: Princeton University Press, 2019.

Watson, Gary. „Responsibility and the limits of evil: Variations on a strawsonian theme“. In *Responsibility, Character, and the Emotions: New Essays in Moral Psychology,* herausgegebene von Ferdinand Schoeman, 256–86. Harvard: Cambridge University Press, 1987.
———. „Promises, reasons and normative power“. In *Reasons for Action*, herausgegeben von David Sobel und Steven Wall, 155–78. Cambridge: Cambridge University Press, 2009.

Williams, Bernard. „Internal and external reasons.“ In *Rational Action*, herausgegeben von Ross Harrison, 101–13. Cambridge: Cambridge University Press, 1979.
———. *Truth & Truthfulness: An Essay in Genealogy*. Princeton: Princeton University Press, 2002.

Williamson, Timothy. *Knowledge and its Limits*. Oxford/New York: Oxford University Press on Demand, 2002.

Wittgenstein, Ludwig. *The Blue and the Brown Books: Preliminary Studies for the Philosophical Investigations.* 2. Auflage. Malden/Oxford/Carlton: Blackwell Publishing, 1969.
———. *Philosophische Untersuchungen.* Herausgegeben von Joachim Schulte. Frankfurt a.M.: Suhrkamp, 2003a.
———. „Unsere Methode.“ In *The Voices of Wittgenstein*, herausgegeben von Gordon Baker, 276-311. London: Routledge, 2003b.
———. *Über Gewißheit*. Herausgegeben von Gertrude Elizabeth Margaret Anscombe und Georg Henrik von Wright. 13. Aufl. Frankfurt a.M.: Suhrkamp, 2015.

Zagzebski, Linda Trinkaus. *Epistemic Authority: A Theory of Trust, Authority, and Autonomy in Belief*. Oxford, New York: Oxford University Press, 2015.

Zylberman, Ariel. „The very thought of (wronging) you“. *Philosophical Topics* 42, Nr. 1 (2014): 153–75.

Lebenslauf des Autors

Persönliche Daten

Name:	Jörn Wiengarn
Geburtsdatum:	20.05.1990

Bildungsgang

01-03/2020	Forschungsaufenthalt an der University of Warwick
Seit 04/2018:	Promotionsstipendiat an der a.r.t.e.s. Graduate School for the Humanities Cologne
09/2017-03/2018:	Predoc-Stipendiat an der a.r.t.e.s. Graduate School for the Humanities Cologne
11/2016:	*Erstes Staatsexamen* in den Fächern *Philosophie/Ethik* und *Politik und Wirtschaft*
2011-2016	Studium auf *LaG* mit den Fächern *Physik*, *Philosophie/Ethik* und *Politik und Wirtschaft* an der TU Darmstadt
2009-2011	Bachelor-Studium *Physik* an der TU Darmstadt
05/2009	Abitur am Gymnasium Georgianum Vreden

Akademische Anstellungen

Seit 02/2021	Wissenschaftlicher Mitarbeiter (Postdoctoral Research Fellow), Philosophie, an der Universität Stuttgart

Vorträge/Präsentationen

„Rationality and agent-specificity in trusting relationships“, Poster-Präsentation, 16.2.2018, Zweiter Workshop „Understanding Others – Integration of Social Cognitive and Affective

Processes“, Universität zu Köln.

„Interpersonal Trust: How to Reconcile Epistemology and Ethics?“, 25.6.2019, II FINO Graduate Conference. Contemporary Issues across Ethics and Epistemology, Universität Pavia.

„Three Concepts of Interpersonal Trust“, 19.09.2019, AHRC DTP Conference on “Trust and Truth”, University of Cambridge.

„Ich habe dir vertraut! - Zur spezifischen Normativität interpersonellen Vertrauens“, 2.10.2019, VII. Tagung für praktische Philosophie, Universität Salzburg.

„Trustworthy Persons, Trustworthy Institutions”, 7.09.2021. MANCEPT Workshop, Manchester Centre for Political Theory.

Lehrerfahrungen

WS 2015-SS 2017:	Konzeption und Lehre von 4 Tutorien am AStA der TU Darmstadt
SS 2012-WS 2017:	Lehre von 6 Tutorien am Fachbereich Physik der TU Darmstadt
SS 2015-WS 2016:	Planung und Lehre von 2 Tutorien am Institut für Philosophie der TU Darmstadt
SS 2020	Seminar „Moralische Gefühle” (zusammen mit Prof. Dr. Christine Chwaszcza)