

Beim folgenden Kapitel handelt es sich um eine pre-published Version der folgenden Veröffentlichung:

Wyler, H. (2021). Befragungsmethoden und Erkenntnisgewinn im Strafverfahren aus psychologischer Sicht. In R. Deckers & G. Köhnken (Hrsg.), *Die Erhebung von Zeugenaussagen im Strafprozess* (Band 4) (S. 110–162). Berlin, Deutschland: Berliner Wissenschafts-Verlag.

<https://www.bwv-verlag.de/detailview?no=5053>

Befragungsmethoden und Erkenntnisgewinn im Strafverfahren aus psychologischer Sicht¹

Helen Wyler

1. Einleitung

Die Vernehmung von Tatverdächtigen ist im Rahmen der Strafverfolgung ein wichtiger Bestandteil der Wahrheitsfindung. Dabei kann die Art und Weise, *wie* eine Vernehmung durchgeführt wird, den Erkenntnisgewinn maßgeblich beeinflussen. Dieser Beitrag soll dem Leser² einen Überblick über den aktuellen Stand der rechtspsychologischen Forschung zum Thema „Befragungsmethoden“ geben, wobei der Fokus auf der Vernehmung von Tatverdächtigen liegt, welche grundsätzlich gewillt sind, eine Aussage zu machen. Innerhalb der Befragungsmethode kann man den übergeordneten Befragungsstil von den spezifischen Befragungsstrategien unterscheiden. Der Befragungsstil beschreibt die allgemeine Herangehensweise an eine Vernehmung, die entweder als konfrontativ-geständnisorientiert oder informationssammelnd klassifiziert werden kann. Befragungsstrategien hingegen sind spezifische Techniken, die innerhalb eines Befragungsstils eingesetzt werden können. Beide verfolgen das Ziel, den Erkenntnisgewinn in einer Vernehmung zu maximieren. So können sich Befragungsstil und -strategie einerseits darauf auswirken, wie viele Informationen die vernommene Person preisgibt. Andererseits kann durch den geschickten Einsatz spezifischer Befragungsstrategien unter Umständen aber auch die Einschätzung erleichtert werden, ob ein Tatverdächtiger die ihm vorgeworfene Tat zu Recht bestreitet oder nicht. Solche Ansätze zur Verbesserung der korrekten Glaubhaftigkeitseinschätzung von Aussagen sind nötig, da wir „intuitiv“ kaum dazu in der Lage sind, Wahrheit von Täuschung³ akkurat zu unterscheiden. Eine Ursache für unsere Schwierigkeiten bei der Lügenerkennung ist darin zu sehen, dass es insgesamt zu wenig Hinweise gibt, die Lüge von Wahrheit zuverlässig unterscheiden. Daher haben sich die Forschungsbemühungen in den letzten 15 Jahren vor allem auf die Erarbeitung von Befragungsstrategien konzentriert, welche die zumeist geringen Unterschiede zwischen lügenden und ehrlichen Tatverdächtigen *vergrößern* und somit das Erkennen von Täuschung erleichtern sollen.

Im Folgenden wenden wir uns zunächst der Frage zu, weshalb die Unterscheidung zwischen wahren und gelogenen Aussagen so schwerfällt. Im Anschluss widmen wir uns dem konfrontativ-geständnisorientierten versus dem informationssammelnden Befragungsstil und stellen ihre jeweiligen Vor- und Nachteile dar. Danach befassen wir uns mit den spezifischen Befragungsstrategien und gehen abschließend der Frage nach, welche Möglichkeiten bestehen, diese Befragungsstrategien in der Praxis anzuwenden und welche Fallstricke es dabei zu beachten gilt.

¹ Mein herzlicher Dank gilt Margit E. Oswald für ihre hilfreichen Anregungen zu diesem Kapitel.

² Aufgrund der besseren Lesbarkeit wird in diesem Text auf die explizite zusätzliche Erwähnung der weiblichen Form verzichtet. Die Autorin weist an dieser Stelle ausdrücklich darauf hin, dass die männliche Form immer als die weibliche Form miteinschließend zu verstehen ist.

³ Die Begriffe „Lüge“ und „Täuschung“ werden synonym verwendet.

2. Probleme und mögliche Lösungsansätze bei der Unterscheidung von Wahrheit und Täuschung

Es ist eine in der wissenschaftlichen Literatur vielfach untersuchte Frage, wie erfolgreich Personen bei der Unterscheidung von Lüge⁴ und Wahrheit sind. In speziell dafür geplanten Studien werden die Teilnehmenden jeweils darum gebeten, eine Anzahl von Aussagen auf ihren Wahrheitsgehalt hin zu beurteilen. Würde das Urteil Wahrheit oder Lüge für jede Aussage durch das Werfen einer Münze gefällt, so erzielte man eine Trefferquote von ca. 50%, da die vorgelegten Aussagen jeweils so zusammengestellt werden, dass die eine Hälfte wahr und die andere Hälfte gelogen ist. Eine Meta-Analyse⁵ von Bond und DePaulo (2006) ergab, dass die Trefferquote von Laien im Schnitt 54% beträgt und somit nur knapp über der durch Zufall zu erwartenden Trefferquote liegt. Sogenannte „lie-catcher“, also Personen, für die das Erkennen von Täuschung im beruflichen Kontext relevant ist, schneiden mit einer durchschnittlichen Trefferquote von 56% nur unwesentlich besser ab (Vrij, 2008a).⁶ Diese Befunde machen deutlich, dass wir kaum dazu in der Lage sind, „intuitiv“ Wahrheit von Täuschung treffsicher zu unterscheiden. Eine sowohl praktisch als auch theoretisch bedeutsame Frage ist es daher, wie die Urteilskorrektheit verbessert werden kann. Hierzu müssen wir jedoch zunächst der *Ursache* der mangelhaften Lügendetektionsleistung auf den Grund gehen.

2.1. Fehlende valide Hinweise versus falsche Nutzung von Hinweisen

Zwei Ansatzpunkte für die Erklärung der niedrigen Trefferquote sind zu unterscheiden (siehe z.B. Hartwig & Bond, 2011). Einerseits ist denkbar, dass Personen im Allgemeinen sehr gut lügen und sich daher nicht oder kaum von Wahrheitsagenden unterscheiden. In diesem Fall werden von den lügenden Personen schlicht zu wenige Hinweise preisgegeben, wodurch es grundsätzlich nicht möglich ist, ihre Lügen korrekt zu identifizieren. Es herrscht in den zu beurteilenden Aussagen also ein *Mangel an Hinweisen* vor, die dazu beitragen könnten, eine korrekte Diagnose von wahr oder falsch zu stellen. (Solche Hinweise werden auch als „valide“ bezeichnet, wie gleich noch näher erläutert wird.) Andererseits könnte das Problem

⁴ Eine Lüge setzt voraus, dass bei der aussagenden Person eine Täuschungsabsicht vorliegt (Vrij, 2008a). Aussagen, die aufgrund von Erinnerungsfehlern nicht den tatsächlichen Geschehnissen entsprechen, werden daher beispielsweise nicht als Lüge klassifiziert.

⁵ Meta-Analysen sind Übersichtsarbeiten, welche die Ergebnisse über viele Studien hinweg zusammenfassen und die durchschnittlich beobachteten Effekte berichten.

⁶ Die Befunde zu den Trefferquoten beziehen sich in ihrer großen Mehrheit auf Laborstudien. Die zu beurteilenden Aussagen stammen dabei von Personen, welche im Rahmen einer Studienteilnahme instruiert wurden, entweder die Wahrheit zu sagen oder aber zu lügen. Ihre Aussagen beziehen sich je nach Studie z.B. auf ein gestelltes Verbrechen („mock crime“), einen beobachteten Sachverhalt, ihre Einstellung zu einem Thema oder ihre Emotionen. Laborstudien haben den zentralen Vorteil, dass der Realitätsgehalt der Aussagen, die sogenannte Grundwahrheit, eindeutig eruiert werden kann. Hingegen steht für die aussagende Person im Labor natürlich wesentlich weniger auf dem Spiel als bei einer richtigen Vernehmung. In ihrer Meta-Analyse fanden Hartwig und Bond (2014) jedoch, dass der Wahrheitsgehalt von Aussagen für stark vs. wenig motivierte Personen anhand der in den Aussagen beobachteten Hinweise ähnlich gut vorhergesagt werden konnte. Auch der Inhalt der Lüge, also ob es z.B. um Fakten oder Emotionen geht, machte keinen Unterschied, was Bedenken bezüglich der Übertragbarkeit der Ergebnisse aus dem Labor auf die Praxis etwas entschärfen kann. Bei Feldstudien werden zwar Aussagen aus echten Polizeibefragungen verwendet, hier kann die Grundwahrheit allerdings oftmals nicht eindeutig festgestellt werden (siehe z.B. Hartwig, 2011). Ferner ist die Beschaffung solcher Daten auch aus Datenschutzgründen eine Herausforderung.

aber auch an der *mangelnden Fähigkeit der Beurteiler* liegen. In diesem Fall wäre selbst dann, wenn hinreichend viele valide Hinweise in den zu beurteilenden Aussagen vorhanden wären, eine schlechte Lügenerkennung zu erwarten, weil die Beurteiler sie entweder nicht erkennen und demzufolge auch nicht nutzen können, oder weil sie sich durch falsche, bzw. diagnostisch irrelevante Hinweise in die Irre führen lassen.

Eine kaum noch überschaubare Anzahl von Studien ist bereits der Frage nachgegangen, ob es denn überhaupt hinreichend viele Hinweise gibt, anhand derer Wahrheit von Täuschung einigermaßen zuverlässig unterschieden werden kann. In diesen Studien müssen zunächst die wahren und gelogenen Aussagen hinsichtlich des Vorliegens oder der Ausprägung bestimmter Merkmale, wie beispielsweise Detailreichtum oder Blickvermeidung, systematisch kodiert werden. Anschließend wird analysiert, ob ein statistisch signifikanter, also grösser als rein durch Zufall zu erwartender Zusammenhang zwischen dem Wahrheitsgehalt der Aussagen (wahr oder falsch) und dem Vorliegen eines bestimmten Hinweises besteht, d.h., ob der Hinweis merkbar häufiger bei einer Lüge als bei einer wahren Aussage vorliegt (oder umgekehrt). Ist dies der Fall, so bezeichnet man diesen Hinweis als „valide“ oder „diagnostisch“. Nun muss man sich aber darüber im Klaren sein, dass statistische Signifikanz nicht automatisch mit praktischer Relevanz gleichgesetzt werden kann. Statistische Signifikanz sagt nichts über das *Ausmaß* des beobachteten Zusammenhangs, die sogenannte Effektstärke, aus. Cohens *d* ist ein gebräuchliches Effektstärkenmaß, bei dem kleine (0.20), mittlere (0.50) und große (0.80) Effekte unterschieden werden (Cohen, 1988). Je grösser die Effektstärke, desto zuverlässiger ist ein Hinweis in der Unterscheidung von Wahrheit und Täuschung. Ein valider Hinweis, der aber nur eine Effektstärke von 0.20 aufweist, ist also für die Praxis nur von begrenztem Nutzen. Es ist daher wichtig, Studien zur Validität von Hinweisen nicht nur hinsichtlich der statistischen Signifikanz der Ergebnisse („Kann man sagen, dass das Ergebnis höchstwahrscheinlich nicht durch Zufall bedingt ist?“), sondern auch in puncto Effektstärke („Wie stark hängt der Hinweis mit dem Wahrheitsgehalt der Aussage zusammen?“) zu studieren.

DePaulo und Kollegen (2003) haben in ihrer Meta-Analyse die Ergebnisse von 120 Studien ausgewertet, welche die Validität verschiedener Hinweise untersucht haben. Von den insgesamt 158 untersuchten Hinweisen standen lediglich 41 in einem statistisch signifikanten Zusammenhang mit dem Wahrheitsgehalt der Aussagen. Davon waren 21 Hinweise in mehr als drei Studien getestet worden, wodurch die Ergebnisse aussagekräftiger sind als für Hinweise, welche nur vereinzelt in Studien eingeschlossen wurden. Wie Tabelle 1 entnommen werden kann, liegen die Effektstärken dieser 21 validen Hinweise mehrheitlich im kleinen bis mittleren Bereich. Insgesamt haben sich inhaltsbezogene, also verbale Hinweise für die Unterscheidung von Wahrheit und Täuschung als nützlicher erwiesen als nonverbale, also körpersprachliche Hinweise (siehe z.B. auch Vrij, 2008a und 2008b). Interessanterweise stimmt dieser Befund so gar nicht mit unseren Vermutungen über die Validität von Hinweisen überein. Sowohl Laien (Global Deception Research Team, 2006) als auch „lie-catcher“ (z.B. Akehurst et al., 1996; Strömwall & Granhag, 2003) berichten im Allgemeinen, dass nonverbale Hinweise, insbesondere Blickvermeidung, Nesteln und sonstiges vermehrtes Bewegen, ihrer Ansicht nach gute Lügenhinweise sind. In Wirklichkeit stehen diese von den Beurteilern berichteten Hinweise aber in keinem oder nur sehr geringfügigem Zusammenhang mit dem Wahrheitsgehalt von Aussagen (DePaulo et al., 2003).

Tabelle 1: Übersicht über die validen Hinweise aus der Meta-Analyse von DePaulo et al. (2003)

| Hinweis | Anzahl Studien | Cohens <i>d</i> |
|--|----------------|-----------------|
| Übergeordnete Hinweise oder Eindrücke | | |
| Verbale und vokale Unmittelbarkeit (Eindruck) ^a | 7 | -0.55 |
| Diskrepanz, ambivalent ^b | 7 | 0.34 |
| Verbale und vokale Unsicherheit (Eindruck) ^c | 10 | 0.30 |
| Nervös, angespannt (Eindruck) | 16 | 0.27 |
| Verbale und vokale Beteiligung ^d | 7 | -0.21 |
| Verbale Hinweise | | |
| Zugeben von Erinnerungslücken | 5 | -0.42 |
| Details | 24 | -0.30 |
| Spontane Selbstverbesserungen | 5 | -0.29 |
| Logische Struktur ^e | 6 | -0.25 |
| Plausibilität | 9 | -0.23 |
| Unmittelbare Wiederholungen ^f | 4 | 0.21 |
| Negative Aussagen und Beschwerden | 9 | 0.21 |
| Paraverbale Hinweise | | |
| Sprechdauer | 4 | -0.35 |
| Stimmliche Anspannung | 10 | 0.26 |
| Stimmhöhe | 12 | 0.21 |
| Nonverbale Hinweise | | |
| Pupillenerweiterung | 4 | 0.39 |
| Kinn heben | 4 | 0.25 |
| Lippen zusammenpressen | 4 | 0.16 |
| Nesteln (Objekt, Körper, Gesicht) | 14 | 0.16 |
| Illustratoren ^g | 16 | -0.14 |
| Angenehmer Gesichtsausdruck ^h | 13 | -0.12 |

Anmerkungen: Es sind nur Hinweise aufgeführt, die in mindestens vier Studien untersucht wurden. Ein negatives Vorzeichen beim Effektstärkenmaß (Cohens *d*) bedeutet, dass der Hinweis häufiger bei wahren als bei gelogenen Aussagen auftritt (Wahrheitshinweis). Kein Vorzeichen bedeutet, dass der Hinweis häufiger bei gelogenen als bei wahren Aussagen auftritt (Lügenhinweis). ^aSubjektiver Eindruck, dass der Sprecher in seiner Aussage direkt, relevant, klar und persönlich herüberkommt ^bDie Mitteilungen des Sprechers scheinen inkonsistent oder die Informationen aus verschiedenen Quellen (z.B. Gesicht und Stimme) scheinen widersprüchlich ^cDer Sprecher scheint unsicher oder nicht sehr bestimmt oder er hat Schwierigkeiten beim Beantworten der Fragen ^dDer Sprecher schildert Ereignisse persönlich und offenherzig, scheint vokal ausdrucksstark und beteiligt ^eKonsistenz und Kohärenz einer Aussage ^fWörter oder Phrasen, die ohne Sprechpause wiederholt werden ^gHandbewegungen, welche das Gesagte begleiten und illustrieren ^hDer Gesichtsausdruck des Sprechers wird als angenehm wahrgenommen, insgesamt mehr positive als negative Gesichtsausdrücke

Zusammenfassend kann also gesagt werden, dass für die wenigen Hinweise, die sich in der Meta-Analyse von DePaulo und Kollegen (2003) als valide herausstellten, die Effekte oft nur klein sind, weshalb sie für den Alltag der Lügendetektion wenig Relevanz besitzen. Von einer sogenannten Pinocchio-Nase kann daher keine Rede sein. Zugleich berichten sowohl Laien als auch „lie-catcher“ mehrheitlich solche Merkmale bei der Lügenerkennung zu benutzen, die nachweislich invalide sind. Es bleibt also die Ausgangsfrage nach dem Hauptgrund für die schlechte Lügendetektionsleistung – ist es der a priori vorhandene Mangel an validen Hinweisen oder die falsche Nutzung von Hinweisen? Hartwig und Bond (2011) haben in ihrer Meta-Analyse anhand statistischer Modelle die beiden Erklärungsansätze gegeneinander getestet. Obwohl Beurteiler durchaus Fehler machen (siehe 2.2.) und nicht alle vorhandenen Hinweise optimal nutzen, scheinen sie doch recht gut darin zu sein, die wenigen einigermaßen validen Hinweise *intuitiv* richtig zu nutzen. Dieses Ergebnis mag überraschen, denn gerade wurde ja gesagt, dass Personen berichten, solche Merkmale bei ihren Urteilen zu verwenden, die nachweislich invalide sind. Allerdings handelt es sich bei diesem Befund um Selbstberichte, also um Angaben darüber, welche Hinweise Personen zu benutzen *glauben*. Gemäß den Analysen von Hartwig und Bond (2011) stimmt diese selbstberichtete Nutzung nur begrenzt mit unserer tatsächlichen Nutzung von Hinweisen überein.⁷ Oder anders gesagt: Wir nutzen andere Hinweise, als die, die wir zu nutzen meinen. Diese Diskrepanz ist darauf zurückzuführen, dass es uns oftmals schwerfällt, die tatsächlichen Beweggründe für unsere Entscheidungen oder Urteile zu identifizieren (vgl. z.B. Nisbett & Wilson, 1977); Psychologen sprechen in diesem Zusammenhang von einer mangelnden Introspektionsfähigkeit. Das Grundproblem der schlechten Lügendetektion liegt also im Mangel an validen Hinweisen, so die zentrale Erkenntnis von Hartwig und Bond (2011). Unter normalen Umständen sind schlicht nicht genügend Hinweise vorhanden, anhand derer Lüge und Wahrheit zuverlässig unterschieden werden könnten.

Die Erkenntnis, dass die Lügenerkennung in erster Linie durch einen Mangel an validen Hinweisen erschwert wird, hat die Forschung dazu inspiriert, spezifischen Befragungsstrategien zu entwickeln, mit denen die zumeist schwachen Unterschiede zwischen Lüge und Wahrheit vergrößert werden können. Dabei geht es primär um die Verstärkung der *verbalen* Unterschiede zwischen lügenden und den die Wahrheit sagenden Personen, da verbale Hinweise, wie bereits erwähnt, generell valider sind. Auf diese neu entwickelten Befragungsstrategien werden wir in Kapitel 4 näher eingehen. Allerdings bedeuten die Befunde von Hartwig und Bond (2011) *nicht*, dass Beurteilern gar keine Urteilsfehler unterlaufen. Zwei zentrale Urteilsfehler, die „Vernachlässigung der Baseline“ und der „Bestätigungsfehler“, sollen daher im Folgenden besprochen werden. Diese beiden Urteilsfehler sind auch beim Einsatz von spezifischen Befragungsstrategien in

⁷ Um die „tatsächliche“ Nutzung von Hinweisen zu erfassen, wird untersucht, ob die von einer Person gefällten Wahrheits- und Lügenurteile systematisch mit dem Auftreten bestimmter Hinweise zusammenhängen. Wenn eine Person also beispielsweise häufig bei Aussagen mit vielen Details ein „wahr“-Urteil abgibt und gleichzeitig Aussagen mit wenigen Details häufig als „gelogen“ einschätzt, wird angenommen, dass die Menge an Details das Urteil der Person beeinflussen könnte. Selbstverständlich handelt es sich hierbei lediglich um Aussagen über beobachtete Zusammenhänge. Daraus kann nicht mit Sicherheit geschlossen werden, dass eine Kausalität vorliegt, also das Vorhandensein von Details eine *Ursache* des „wahr“-Urteils war. Hingegen würde man aber durchaus erwarten, dass sich ein starker Zusammenhang zwischen z.B. Urteil und Blickvermeidung zeigen würde, wenn Personen den Hinweis „Blickvermeidung“ systematisch bei ihrem Urteil berücksichtigen würden.

Vernehmungen von zentraler Bedeutung. Wir werden sie daher gegen Ende dieses Beitrags nochmals antreffen.

2.2. Wichtige Urteilsfehler

Urteilsfehler 1: Vernachlässigung der Baseline

Nehmen wir einmal an, wir möchten den Wahrheitshinweis „Detailreichtum“ zur Beurteilung einer Aussage nutzen. Gehen wir ferner davon aus, dass extravertierte Personen gesprächiger sind als introvertierte Personen und somit auch detaillierter berichten. Wenn unser Urteil nun *ausschließlich* auf der Aussage abstellt, welche eine Person zum kritischen Ereignis macht, könnte es sein, dass wir Aussagen von extravertierten Personen häufig als wahr einschätzen und Aussagen von introvertierten Personen häufig als gelogen klassifizieren, und zwar unabhängig vom tatsächlichen Wahrheitsgehalt der Aussage. Das wäre natürlich nicht wünschenswert. Diesem Problem können wir entgegenwirken, indem wir eine „Baseline“ (im Sinne einer „Vergleichswahrheit“) erheben, d.h. berücksichtigen, wie sich die Person normalerweise verhält, wenn sie die Wahrheit sagt. Anhand der Baseline kann dann eingeschätzt werden, ob es zu Veränderungen im Verhalten einer Person kommt, wenn diese zum kritischen Ereignis aussagt (z.B. die Ausführungen sind weniger detailliert als zuvor).

Viele Befrager haben die Relevanz der Baseline zwar grundsätzlich erkannt, allerdings wird sie häufig nicht korrekt erhoben (siehe z.B. Vrij, 2016). So werden in Vernehmungen nicht selten Small Talk oder das Beantworten von nicht-tatrelevanten Fragen zu Beginn der Vernehmung als Baseline verwendet (Vrij, 2016). Mit diesem Vorgehen wird jedoch der Einfluss vernachlässigt, den die Situation auf das Verhalten einer Person hat, denn die Situation beim Small Talk oder beim Beantworten nicht-tatrelevanter Fragen unterscheidet sich in vielerlei Hinsicht von jener in der eigentlichen Vernehmung. Unsere Tendenz, das Verhalten einer Person (z.B. Nervosität) auf Faktoren innerhalb der Person zurückzuführen (Person lügt) und dabei den Einfluss der Situation oder der Umstände zu vernachlässigen (Person befindet sich in einer ungewohnten, stressreichen Situation und hat Angst, dass ihr fälschlicherweise nicht geglaubt wird), bezeichnen Psychologen als „fundamentalen Attributionsfehler“ (siehe z.B. Fincham & Hewstone, 2003). Gemäß einer Laborstudie verändert beispielsweise schon nur das bloße Wissen darum, dass der Wahrheitsgehalt der Aussage nun beurteilt wird, das nonverbale Verhalten – und zwar auch bei Personen, die die Wahrheit sagen (Palena et al., 2018). Darüber hinaus steht natürlich sowohl für die Wahrheit sagende als auch für lügende Personen während der eigentlichen Befragung viel mehr auf dem Spiel als während der erhobenen Baseline. Daher sind Small Talk oder nicht-tatrelevante Fragen zu Beginn der Befragung als Baseline ungeeignet.

Da sich die Vergleichssituation und die zu beurteilende Situation so ähnlich wie möglich sein müssen, kann eine angemessene Baseline daher nur *in der kritischen Befragungssituation selbst* erhoben werden (Vrij, 2016). Hier ist der Vernehmer aber mit dem Problem konfrontiert, dass er nicht mit Sicherheit weiß, ob die Person in der Vergleichssituation tatsächlich die Wahrheit sagt, also wirklich eine „Vergleichswahrheit“ vorliegt. Das Etablieren einer angemessenen Baseline ist daher ausgesprochen schwierig. Einige der neuen Befragungsstrategien haben nun den Vorteil, dass sie anderweitig Vergleiche innerhalb der Person zulassen, wodurch die eingangs erwähnten Unterschiede zwischen den

Personen berücksichtigt werden können, ohne dass eine Vergleichswahrheit im eigentlichen Sinne erhoben werden muss. Auf diesen Punkt kommen wir später nochmals zurück.

Urteilsfehler 2: Bestätigungsfehler

Der Bestätigungsfehler beschreibt unsere Tendenz, eine einmal getroffene Annahme oder Erwartung beibehalten zu wollen (siehe z.B. Oswald & Wyler, 2018). Dazu werden Informationen systematisch in einer unsere Annahme bestätigenden Art und Weise selektiv gesucht, verarbeitet und erinnert. Die Vernehmungssituation bietet ideale Bedingungen für das Auftreten eines Bestätigungsfehlers. Volbert und May (2016) sprechen in diesem Zusammenhang von einer „strukturbedingten Schuldverzerrung“. Der Vernehmung einer Person als Tatverdächtiger gehen Ermittlungen voraus, welche gewisse belastende Informationen hervorgebracht haben. Dadurch ist es wahrscheinlich, dass der Ermittler – selbst bei den besten Absichten – mit einer gewissen Schuldvorannahme in die Befragung hineingeht. Diese Schuldvorannahme kann dann u.a. dazu führen, dass primär (vermeintliche) Lügenhinweise wahrgenommen werden, während allfällige Wahrheitshinweise weniger erkannt oder beachtet werden. Aber auch an den Vorermittlungen unbeteiligte Befragte sind vor einem Bestätigungsfehler nicht gefeit. So zeigen Studien, dass der erste Eindruck einer Person blitzschnell gebildet wird, z.B. hinsichtlich ihrer Vertrauenswürdigkeit (Porter et al., 2008). Auch dieser erste Eindruck kann sich dann auf die weitere Informationsverarbeitung auswirken (Porter et al., 2010).

Dem Bestätigungsfehler kann am ehesten entgegengewirkt werden, indem stetig auch alternative Erklärungsmöglichkeiten für gemachte Beobachtungen und vorhandene Beweise formuliert und geprüft werden (siehe z.B. Oswald & Grosjean, 2004; Rassin, 2010). Dieses Prüfen von Alternativhypothesen könnte besonders dann hilfreich sein – und auch etwas leichter fallen – wenn die Situation mit einer gewissen Distanz evaluiert werden kann, z.B. während Vernehmungspausen oder im Anschluss an eine Vernehmung. Dieses Prüfen von Alternativerklärungen ist auch beim Einsatz von spezifischen Befragungsstrategien äußerst wichtig, um das Risiko von Fehleinschätzungen möglichst gering zu halten. Darauf werden wir später noch näher eingehen.

Fazit

Es hat sich anhand zahlreicher Studien zusammenfassend gezeigt, dass es in Vernehmungssituationen äußerst schwierig ist, Lüge von Wahrheit zu unterscheiden. Dies vor allem dann, wenn sich das Urteil ausschließlich auf die Aussagen der vernommenen Personen stützen muss. Lügende Personen unterscheiden sich von denen, die die Wahrheit sagen, nur durch wenige valide Hinweise, die aber mehrheitlich schwach ausgeprägt sind. Schwach sind diese Hinweise deshalb, weil sie keineswegs immer oder zumindest meistens, sondern nur etwas öfter bei Lüge als bei Wahrheit vorkommen (oder umgekehrt). Interessanterweise scheint der Beurteiler diese wenigen validen Hinweise aber intuitiv in seinem Urteil recht gut zu berücksichtigen. Dennoch kann er systematische Fehler begehen. So kann er die wahrgenommenen Hinweise nicht hinreichend in Bezug auf die spezielle Vernehmungssituation oder die Persönlichkeit der vernommenen Person relativieren (Vernachlässigung der Baseline), oder sich von einer einmal gebildeten Annahme über den Wahrheitsgehalt einer Aussage zu lange leiten lassen (Bestätigungsfehler).

3. Übergeordneter Ansatz: Informationssammelnder vs. konfrontativ-geständnisorientierter Befragungsstil

Betrachten wir nun zunächst den allgemeinen, übergeordneten Ansatz, mit dem an eine Vernehmung herangegangen werden kann. Wie erwähnt können wir hier im Wesentlichen zwei Stile oder Vernehmungstechniken unterscheiden, nämlich den konfrontativ-geständnisorientierten Befragungsstil und den informationssammelnden Befragungsstil (Moston & Engelberg, 1993; siehe auch Vrij, Meissner et al. 2017). Dabei handelt es sich um zwei grundverschiedene Arten, eine Vernehmung anzugehen: Der eine Befragungsstil hat das Erlangen eines Geständnisses zum Ziel, während beim anderen die Informationsgewinnung im Vordergrund steht. Die beiden Stile sollen im Folgenden erörtert und evaluiert werden.

3.1. Konfrontativ-geständnisorientierter Befragungsstil und die Reid-Technik

Der konfrontativ-geständnisorientierte Befragungsstil zeichnet sich dadurch aus, dass der Befrager von der Schuld des Befragten ausgeht und daher in der Befragung das Ziel verfolgt, vom Befragten ein Geständnis zu erlangen (siehe z.B. Meissner et al., 2014 und Vrij, Meissner et al., 2017). Der Befrager kommuniziert dem Befragten offen, dass er ihn für schuldig hält. Oftmals wird der Vernommene früh in der Befragung mit den gegen ihn vorliegenden Beweisen konfrontiert, und es werden diverse psychologisch-manipulative Taktiken eingesetzt, um ein Geständnis zu erzielen.⁸ Die Kontrolle des Gesprächs liegt beim Befrager, welcher mehrheitlich geschlossene Fragen stellt, die darauf abzielen, seine eigenen Annahmen zu den Ereignissen zu bestätigen. Durch dieses Vorgehen wird der Befragte in die Defensive gedrängt und seine Antworten konzentrieren sich oftmals auf das Zurückweisen der gegen ihn vorgebrachten Anschuldigungen (Vrij et al., 2007).

Die insbesondere in Nordamerika verbreitete Reid-Methode (Inbau et al., 2013) ist die Versinnbildlichung eines konfrontativ-geständnisorientierten Befragungsstils. Die Reid-Methode besteht aus zwei Phasen. In der ersten Phase führt der Befrager zunächst eine neutrale, nicht anschuldigende Vorbefragung durch, welche als „Behavioral Analysis Interview“ oder „Verhaltens-Analyse-Interview“ bekannt ist. Kommt der Befrager nach seiner Analyse des allgemeinen nonverbalen Verhaltens des Befragten sowie dessen Antworten auf bestimmte Fragen zum Schluss, die Schuld des Befragten scheinbar „definite or reasonably certain“ (Inbau et al., 2013, S. 187), also zweifelsfrei oder zumindest relativ sicher, folgt in einer zweiten Phase eine konfrontativ-geständnisorientierte neunstufige Befragung, welche den Aussagewiderstand des (vermeintlichen) Täters brechen soll (Inbau et al., 2013; siehe auch Brockmann, 2008, S. 249). In dieser zweiten Phase werden gezielt verschiedene Arten psychologischer Manipulation eingesetzt, um ein Geständnis zu erreichen. Laut Kassin und Gudjonsson (2004, S. 43) können die neun Stufen der Befragung auf das Zusammenspiel der folgenden drei Prozesse reduziert werden: (1) Haft und Isolation, welche zu einem erhöhten Stresslevel führen und Optionen attraktiv erscheinen lassen, die es einem erlauben, sich dieser Situation zu entziehen; (2) Konfrontation, bei der der Befrager seine Überzeugung kundtut, dass der Befragte der Täter ist, auf tatsächliche

⁸ Für eine Übersicht über verschiedene psychologisch-manipulative Taktiken, siehe z.B. Bull und Milne (2004) sowie Kassin und Gudjonsson (2004).

oder erfundene Beweise verweist und Versuche des Verdächtigten vereitelt, die Vorwürfe abzustreiten; (3) Minimierung, bei welcher der Befrager sein Verständnis für das Verbrechen signalisiert und moralische Rechtfertigungen anbietet, was vom Verdächtigten dahingehend aufgefasst werden kann, dass er nachsichtig behandelt werden wird. Das Zusammenspiel dieser drei Faktoren kann dazu führen, dass ein Geständnis als beste, wenn nicht gar einzige Option in der aktuellen Situation wahrgenommen wird. Dies gilt allerdings nicht nur für schuldige Tatverdächtige, sondern auch für unschuldige, wie wir gleich noch sehen werden.

Die Reid-Methode ist aus verschiedenen Gründen äußerst umstritten (siehe z.B. Vrij, Meissner et al., 2017). Ein zentrales Problem ist, dass die nonverbalen und inhaltlichen Merkmale, welche Befrager gemäß dem Reid-Manual für die Beurteilung der Vorbefragung nutzen sollen, einer wissenschaftlichen Überprüfung nicht standhalten (Blair & Kooi, 2004; DePaulo et al., 2003; Vrij et al., 2006; siehe auch Vrij, Meissner et al., 2017). Folglich verlassen sich die Befrager für ihr Urteil auf Hinweise, die sich überhaupt nicht dazu eignen, die relevante Frage – lügt die Person oder streitet sie die Tat zu Recht ab – zu beantworten. Hinzu kommt, dass die auf dem Ergebnis der Vorbefragung begründete Schuldannahme oft mit einer überhöhten Urteilssicherheit einhergehen kann, was nicht zuletzt die Folge der vermeintlichen Wissenschaftlichkeit des Ansatzes ist (siehe Kassin & Fong, 1999). Diese Schuldannahme wiederum kann den weiteren Verlauf der Befragung wesentlich beeinflussen und verzerren (siehe z.B. Hill et al., 2008). Ferner weiß man mittlerweile, dass konfrontativ-geständnisorientierte Befragungen durch das Zusammenspiel der drei im vorangegangenen Abschnitt beschriebenen Prozesse zu falschen Geständnissen führen können (Kassin et al., 2010; Meissner et al., 2014). Daran ändert auch die explizite Behauptung von Inbau und Kollegen (2013, S, 187), dass ihr Neun-Stufen-Ansatz nicht dazu „geeignet“ sei, bei einer unschuldigen Person ein Geständnis hervorzurufen⁹, nichts. Die Situation wird zusätzlich dadurch verschärft, dass es beispielsweise Polizisten gibt, die die Auftretenswahrscheinlichkeit von falschen Geständnissen in ihren Befragungen als sehr gering einschätzen (US-Studie; Kassin et al., 2007).¹⁰ Die Möglichkeit eines falschen Geständnisses wird in der Praxis also womöglich gerade von jenen Personen unterschätzt, welche solch problematische Befragungsstile anwenden. Von den eben diskutierten Problemen der REID-Technik abgesehen sind schließlich diverse der von Inbau und Kollegen propagierten Taktiken, wie beispielsweise den Befragten hinsichtlich des Vorhandenseins von Beweisen zu täuschen, weder in Deutschland noch in der Schweiz mit der geltenden Strafprozessordnung vereinbar (Böhme, 2018; Haas & Ill, 2013).

3.2. Informationssammelnder Befragungsstil und das PEACE-Modell

Laut Miller et al. (2018) wird der informationssammelnde Befragungsstil u.a. in diversen Ländern Europas sowie in Australien und Neuseeland eingesetzt, wobei in der Studie leider keine deutschsprachigen Länder untersucht wurden. Aber beispielsweise auch für

⁹ Konkret schreiben die Autoren: „For those investigators who have qualms or reservations about utilizing some of the steps, our discussion of the interrogation process will include explanations as to why these approaches are necessary to persuade a guilty person to tell the truth *and would not be apt to cause an innocent suspect to confess*“ (Inbau et al., 2013, S. 187; Hervorhebung durch die Autorin).

¹⁰ Für Informationen zur Häufigkeit falscher Geständnisse im deutschsprachigen Raum, siehe z.B. Volbert (2012).

Deutschland legen die Bestimmungen in § 136 der StPO „einen informationssammelnden Ansatz nahe“ (Volbert, 2013, S. 232). Das Ziel des informationssammelnden Ansatzes ist es, durch Informationsgewinnung die „Wahrheit“ aufzudecken. Es sollen also möglichst viele und (soweit möglich) korrekte Informationen von der befragten Person erlangt werden. Um dies zu erreichen, soll der Befrager eine Beziehung zum Befragten aufbauen, aktiv zuhören, und dem Befragten durch das Stellen offener Fragen die Möglichkeit geben, seine Sicht der Ereignisse zu schildern (Vrij, Meissner et al., 2017; Volbert, 2013). Die Haltung des Befragers sollte dabei möglichst offen und unvoreingenommen sein. Auch im Rahmen dieses Befragungstils kann der Befragte mit vorhandenen Beweismitteln oder Unstimmigkeiten innerhalb der Aussage konfrontiert werden (z.B. Meissner et al., 2014). Dies geschieht aber erst, nachdem der Befragte seine Version der Ereignisse schildern konnte. Zudem kann und soll der Befragte zu den offengelegten Beweisen Stellung beziehen. Der Beziehungsaufbau sowie die offene Herangehensweise durch den Befrager führen gemäß Studien dazu, dass sich Beschuldigte eher auf ein Gespräch einlassen und auch kooperativer sind (z.B. Bull & Soukara, 2010; Evans et al., 2013).

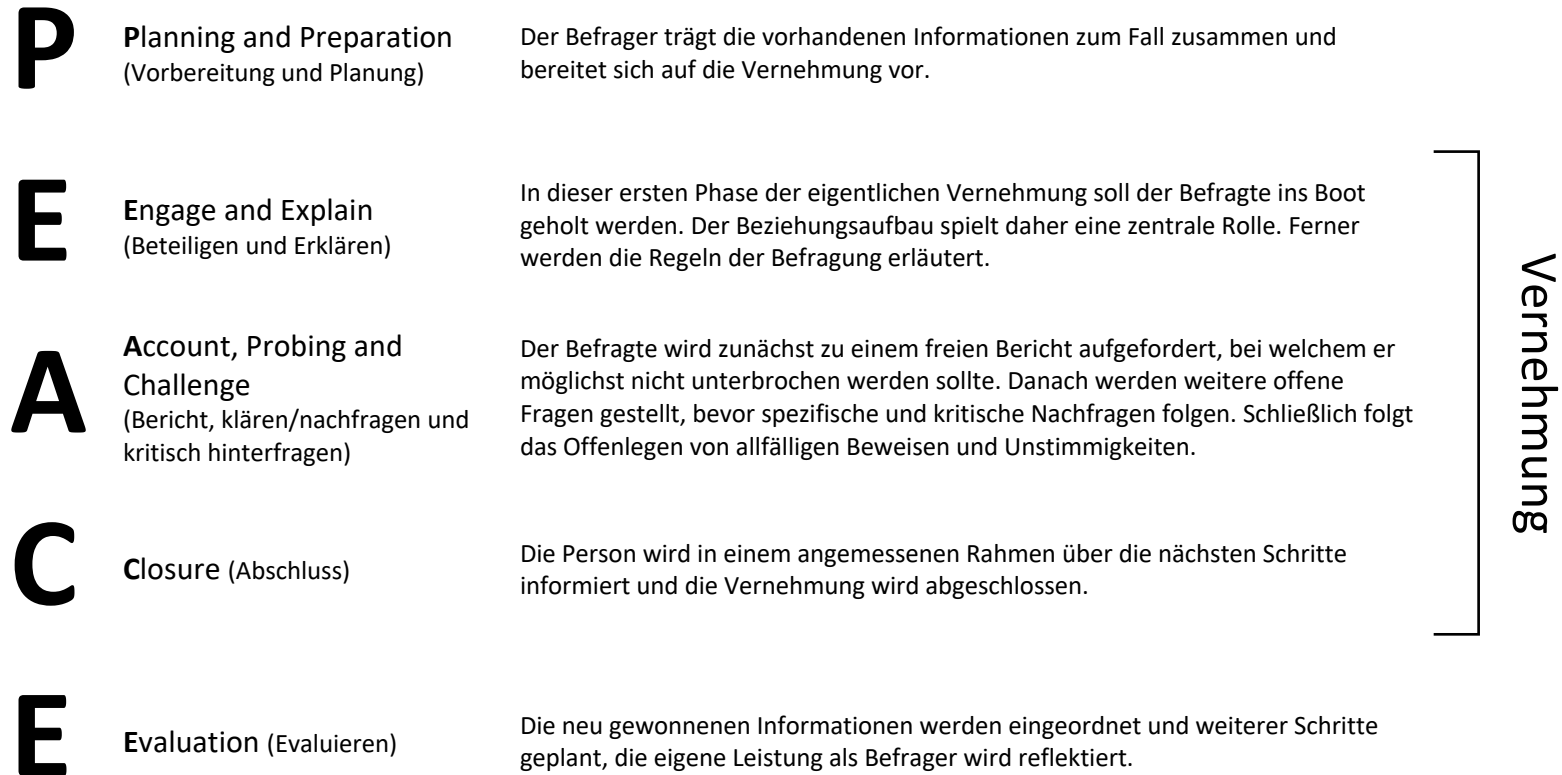
Das PEACE-Modell ist die Verkörperung des informationssammelnden Befragungstils. Es wird in Großbritannien seit 1993 flächendeckend geschult und systematisch bei Vernehmungen eingesetzt (Milne et al., 2007).¹¹ PEACE kommt sowohl bei Tatverdächtigen als auch bei Zeugen zum Einsatz und hat zum Ziel, vom Vernommenen eine möglichst vollständige und möglichst korrekte Aussage zu erhalten. PEACE ist ein Akronym, das für die fünf Schritte einer Befragung steht (siehe Abbildung 1).

Im Rahmen des PEACE-Trainings werden die Polizisten zudem im „kognitiven Interview“ (Fisher & Geiselman, 1992) für aussagewillige Personen und dem „Conversation Management“ (Shepherd, 1993, zit. in Milne et al., 2007) für Aussagende mit geringer Aussagebereitschaft geschult. Die Website des UK College of Policing¹² ist eine hilfreiche Quelle mit weiteren Informationen dazu, welche konkreten Aspekte bei den einzelnen Stufen von PEACE zu berücksichtigen sind (z.B. Frageformulierung, Beziehungsaufbau etc.).

¹¹ Historisch hatte der Wechsel von einem konfrontativ-geständnisorientierten zu einem informationssammelnden Befragungstil seinen Ursprung in einer Reihe spektakulärer Justizirrtümer, bei denen auch problematische Polizeibefragungen eine Rolle spielten (z.B. Guildford Four; Milne et al., 2007). Als Folge dieser Fälle wurde 1984 der Police and Criminal Evidence Act (PACE) verabschiedet, welcher u.a. festlegt, dass Befragte keinem unnötigen polizeilichen Druck oder psychologischer Manipulation ausgesetzt werden dürfen, und dass alle Befragungen auf Polizeistationen aufgezeichnet werden müssen. Eine anschließende Auswertung solcher Aufzeichnungen legte Handlungsbedarf in Bezug auf die Vernehmungspraxis offen, woraufhin erfahrene Polizeibeamte das PEACE-Befragungs-Modell erarbeiteten. Dabei wurden auch wissenschaftlicher Befunde aus der kognitiven Psychologie und der Sozialpsychologie berücksichtigt (Bull & Soukara, 2010; Milne et al., 2007).

¹² <https://www.app.college.police.uk/app-content/investigations/investigative-interviewing/>

Abbildung 1. Übersicht über das PEACE-Befragungs-Modell



3.3. Vergleich der beiden Befragungsstile

Vrij et al. (2007) haben in einem Experiment den informationssammelnden und den konfrontativ-geständnisorientierten Befragungsstil miteinander verglichen. Sie haben festgestellt, dass ersterer zu längeren Antworten führte und diese Antworten auch mehr valide inhaltliche Hinweise enthielten. Ähnlich berichten auch Evans et al. (2013), dass der informationssammelnde Befragungsstil im Vergleich zum konfrontativ-geständnisorientierten Stil zu mehr Kooperation und dem Berichten von mehr relevanten Details führte. Für aussagewillige Personen bietet eine informationssammelnde Befragung zudem bessere Bedingungen für einen erfolgreichen Erinnerungsabruf (siehe Vrij, Meissner, et al., 2017). Die Art des Befragungsstils wirkt sich aber nicht nur auf die aussagende Person aus, sondern beeinflusst auch die Wahrnehmung und Urteile von unbeteiligten Beobachtern. Vrij et al. (2007) fanden, dass Beobachter, welche sich die Befragungen auf Video ansahen, bei konfrontativ-geständnisorientierten Befragungen häufiger fälschlicherweise ehrliche Personen als Lügner einstufte und sich in diesen falschen Urteilen auch besonders sicher waren. Insbesondere bei *unschuldigen* Tatverdächtigen könnte der Einsatz einer konfrontativ-geständnisorientierten Befragung also zu einer weiteren Festigung der bereits vorbestehenden Schuldannahme des Befragers führen. Dies könnte wiederum zur Folge haben, dass mehr Druck auf den Befragten ausgeübt wird (Kassin et al., 2003) und dadurch das Risiko eines falschen Geständnisses steigt. In der Tat erhöhen konfrontative Befragungen im Vergleich mit einer Kontrollbedingung (direkte Befragung) nicht nur die Wahrscheinlichkeit von wahren Geständnissen, sondern auch jene von falschen Geständnissen (Kassin et al., 2010; Meissner et al., 2014). Im Gegensatz dazu führen informationssammelnde Befragungen im Vergleich zur Kontrollbedingung zu einer höheren Anzahl an wahren Geständnissen, *ohne* dabei die Wahrscheinlichkeit von falschen Geständnissen zu erhöhen (Meissner et al., 2014, S. 479). Der informationssammelnde Befragungsstil führt folglich zu diagnostischeren Geständnissen.

Der eine oder andere Leser mag nun vielleicht zu denken geneigt sein, der konfrontativ-geständnisorientierter Befragungsstil betreffe ihn insofern nicht, da solche Vorgehensweisen beispielsweise in Deutschland nicht anzutreffen seien. Die Situation scheint allerdings so einfach nicht zu sein. So fanden in Deutschland – zumindest während einer gewissen Zeit – Schulungen zur Reid-Methode statt (Habschick, 2012; Hermanutz & Litzcke, 2009; Klossek, 2004, zit. in Hermanutz & Litzcke, 2009). In einer Studie gaben von 48 befragten Kriminalpolizisten immerhin 42% an, an einer solchen Schulung teilgenommen zu haben (Litzcke & Klossek, 2009), wobei diese Auswertung für zwei Dienststellen natürlich nicht repräsentativ ist und die Ergebnisse für andere Dienststellen womöglich anders ausfallen. Aber auch sonst kann man bei einer Vernehmung schnell einmal in einer Grauzone landen, insbesondere dann, wenn der Beschuldigte nicht kooperiert. Volbert und May (2016, S. 6) führen dazu aus, dass die deutschsprachige Vernehmungsliteratur zwar keine Taktiken zur Erhöhung der Geständnisbereitschaft vermittele, die Grenzen zwischen der Erhöhung der Aussage- und Geständnisbereitschaft jedoch leicht verschwimmen würden. So gehörten zu den empfohlenen Taktiken zur Erhöhung der „Aussagebereitschaft“ bei unkooperativen Beschuldigten etwa der Hinweis auf die mögliche strafmildernde Wirkung eines Geständnisses oder die Empfehlung, bei vermuteter Scham aufgrund der Tat dem Beschuldigten Verständnis entgegenzubringen. Letztere Taktik gehört zu den Minimierungstechniken, welche auch beim anklagend-geständnisorientierten Befragungsstil

zum Einsatz kommen (Bull & Milne, 2004). Ähnlich kann auch der Hinweis, ein Geständnis sei aus dem einen oder anderen Grund im Interesse des Vernommenen, problematisch sein (siehe z.B. Horgan et al., 2012; Volbert, 2013). Obwohl der konfrontativ-geständnisorientierte Befragungsstil im deutschsprachigen Raum sicher weniger verbreitet ist als beispielsweise in Nordamerika, ist es also dennoch wichtig, sich mit diesem Befragungsstil auseinanderzusetzen und sich der Probleme bewusst zu sein, die mit ihm einhergehen können.

Fazit

In Hinblick auf die Frage, wie der Erkenntnisgewinn im Strafverfahren durch den verwendeten Befragungsstil erhöht werden kann, können wir folgende Vorteile des informationssammelnden Befragungsstils gegenüber dem konfrontativ-geständnisorientierten Befragungsstil festhalten. Der informationssammelnde Ansatz führt zu einer höheren Kooperation der Befragten, fördert mehr relevante Informationen zu Tage und generiert Aussagen, welche mehr valide Glaubhaftigkeitshinweise enthalten, was wiederum bessere Voraussetzungen für eine Beurteilung der Glaubhaftigkeit schafft (zeitnahe durch den Befrager oder zu einem späteren Zeitpunkt). Legt die befragte Person ein Geständnis ab, ist dieses zudem diagnostischer, wenn es im Rahmen einer informationssammelnden Befragung erfolgte. Schließlich bietet der informationssammelnde Befragungsstil aussagewilligen Personen bessere Bedingungen für einen erfolgreichen Erinnerungsabruf. Mit dem PEACE-Modell steht ein gutes Rahmenwerk für informationssammelnde Befragungen zur Verfügung, welches genügend Spielraum lässt, um spezifische Befragungsstrategien einzusetzen, welche dem Befrager oder auch späteren Instanzen helfen können, den Wahrheitsgehalt der getätigten Aussage besser einzuschätzen.

4. Spezifische Vorgehensweisen: Neue Befragungsstrategien

Da die Unterscheidung von Lüge und Wahrheit insbesondere durch einen Mangel an validen Hinweisen erschwert wird, wurden in den vergangenen Jahren diverse Befragungsstrategien entwickelt, welche die Unterschiede zwischen lügenden und die Wahrheit sagenden Personen verstärken sollen. Obwohl sich die verschiedenen Befragungsstrategien in ihrer genauen Funktionsweise unterscheiden, ist ihnen gemein, dass sie Unterschiede in den geistigen Prozessen und Strategien von lügenden und die Wahrheit sagenden Personen nutzen, um zusätzliche valide Hinweise hervorzurufen oder aber bestehende Hinweise zu verstärken. Ein wichtiger Unterschied in den geistigen Prozessen ist, dass Lügen im Normalfall kognitiv belastender ist, als die Wahrheit zu sagen. Lügende und die Wahrheit sagende Personen unterscheiden sich aber auch in den Strategien, die sie während einer Befragung einsetzen, um den Befrager von ihrer Unschuld zu überzeugen (sogenannte „counter-interrogation strategies“). Bevor wir uns den einzelnen Befragungsstrategien zuwenden, gehen wir zunächst auf die Unterschiede in der kognitiven Belastung ein, da diese bei allen Befragungsstrategien zumindest eine gewisse Rolle spielen.

4.1. Die Rolle der kognitiven Belastung

Unsere kognitiven Ressourcen sind begrenzt. Wir alle erleben das in unserem Alltag, wenn wir versuchen, zu viele Dinge gleichzeitig zu erledigen. Irgendwann haben wir keine freien Ressourcen mehr – der Arbeitsspeicher des Rechners ist sozusagen voll ausgelastet – und es passieren uns Fehler. Wie Vrij (2015, S. 206ff) in seinem Übersichtskapitel schön herausgearbeitet hat, ist Lügen aus verschiedenen Gründen kognitiv anspruchsvoller, als die Wahrheit zu sagen. Zu diesen Gründen gehört, dass erstens das Formulieren einer Lüge an sich kognitiv anspruchsvoller ist, als die Wahrheit zu sagen. Die Lüge muss plausibel und möglichst kompatibel mit allfälligen Informationen sein, die dem Befragten bekannt sein könnten. Zudem muss die Lüge vom Befragten später korrekt erinnert werden können, damit er sich nicht selbst widerspricht. Zweitens hält es eine lügende Person für weniger selbstverständlich, dass man ihrer Aussage Glauben schenken wird. Daher neigt sie eher dazu, ihr eigenes Verhalten zu kontrollieren (z.B. nicht nesteln, Augenkontakt halten), was wiederum Ressourcen kostet. Aus dem gleichen Grund überwacht drittens eine lügende Person auch eher die Reaktion ihres Gegenübers auf ihre Lüge. Viertens beansprucht bei einer Lüge auch das aktive Unterdrücken der Wahrheit kognitive Ressourcen. Schließlich ist fünftens das Aktivieren einer Lüge etwas Intentionales und daher mit kognitiver Belastung verbunden, wohingegen die Wahrheit automatisch aktiviert wird. Die Idee, dass die kognitive Belastung beim Lügen zumeist höher ist als bei einer wahren Aussage, ist auch empirisch breit abgestützt (siehe Vrij, 2015, für eine Übersicht). Nun erfordert der Prozess des Erinnerns, der für eine wahre Aussage notwendig ist, natürlich ebenfalls gewisse kognitive Ressourcen. Die Annahme ist jedoch, dass die Belastung *im Regelfall* deutlich geringer ausfällt als bei einer Lüge. Dennoch kann auch eine wahre Aussage mit relativ hohem kognitivem Aufwand verbunden sein, beispielsweise dann, wenn das Ereignis lange her und damit schwieriger zu erinnern ist (siehe z.B. Crozier et al., 2017).¹³

Inwiefern sind diese Annahmen und Erkenntnisse zur kognitiven Belastung nun für die neuen Befragungsstrategien relevant? Vrij (2015) argumentiert, dass das Lügen kognitiv umso anspruchsvoller ist, je stärker die oben erwähnten Punkte zutreffen. Im Umkehrschluss bedeutet das zum Beispiel, dass eine vorbereitete Lüge die Belastung während der Vernehmung reduziert, da der Prozess des Formulierens der Lüge bereits vorweggenommen wurde. Und weiß eine lügende Person, welche Informationen dem Befragten bereits vorliegen, erleichtert dies das Lügen, da sie die Frage, was der Befragte wissen *könnte*, nicht mehr in ihre Überlegungen mit einzubeziehen braucht. Einige der neuen Befragungsstrategien versuchen nun, solchen kognitiven Erleichterungen entgegenzusteuern. So verringert beispielsweise die Strategie des Stellens unerwarteter Fragen die mit dem Vorbereiten einer Lüge verbundenen Vorteile, während die Technik der strategischen Nutzung von Evidenz zurückhaltend mit Informationen darüber ist, was den Ermittlern bereits bekannt ist. Bei einer erhöhten kognitiven Belastung erwartet man ein vermehrtes Auftreten von spezifischen non- und paraverbalen¹⁴ Hinweisen (z.B. mehr Sprechverzögerungen und Sprechfehler oder eine längere Dauer bis zum Einsetzen der

¹³ Mit diesem Problem sehen sich teilweise auch unschuldige Tatverdächtige konfrontiert, wenn sie ein Alibi vorlegen sollen. Die Zeitspanne, für die das Alibi gefragt ist, wurde zum damaligen Zeitpunkt womöglich in keinsten Weise als relevant oder distinkt wahrgenommen, weshalb es umso schwieriger ist, sich nun daran zu erinnern (Crozier et al., 2017).

¹⁴ Paraverbale Hinweise beziehen sich auf die Art und Weise, *wie* etwas gesagt wird.

Antwort; für eine Übersicht, siehe Vrij, 2015), die erhöhte Belastung kann sich aber auch auf verbale Hinweise auswirken, wie bei den einzelnen Befragungsstrategien ersichtlich werden wird. Mit diesen Überlegungen im Hinterkopf können wir nun auf die einzelnen Strategien näher eingehen.

4.2. Unerwartete Fragen stellen

Wenn sie die Gelegenheit dazu haben, bereiten lügende Personen sich oftmals aktiv auf die Vernehmung vor, indem sie ihre Aussagen sowie Antworten auf von ihnen erwartete Fragen einüben (z.B. Hartwig et al., 2007). Diese Vorbereitung erlaubt es ihnen, in der Befragung ihre Antworten aus dem Gedächtnis abzurufen, statt diese an Ort und Stelle erfinden zu müssen. Das Lügen wird also einfacher. Damit übereinstimmend wurde festgestellt, dass vorbereitete Lügen weniger Hinweise enthalten als unvorbereitete (DePaulo et al., 2003), und dass sie auch schwerer zu entdecken sind (Bond & DePaulo, 2006). Das Vorbereiten von Antworten auf spezifische Fragen ist aber nur dann von Nutzen, wenn die lügende Person die Fragen richtig antizipiert hat. Somit können Befrager durch das Stellen unerwarteter Fragen die Strategie der lügenden Person untergraben. Natürlich können lügende Personen eine Antwort auf eine unerwartete Frage verweigern oder aber sagen, sie könnten sich nicht erinnern. Da sie aber vermutlich davon ausgehen, dass ein solches Verhalten Misstrauen hervorruft, dürften sie motiviert sein, die Fragen so gut als möglich aus dem Stehgreif zu beantworten (siehe z.B. Vrij, 2015). Im Folgenden sollen nun drei Arten unerwarteter Fragen besprochen werden.

Rückwärts erzählen

Wenn wir ein Ereignis schildern, tun wir dies zumeist in (einigermaßen) chronologischer Reihenfolge. Wird eine Person nun dazu aufgefordert, das Ereignis in umgekehrter Reihenfolge, also vom Ende hin zum Anfang zu erzählen, erhöht das ihre kognitive Belastung aus drei Gründen. Erstens läuft das Rückwärtserzählen entgegen der natürlichen Kodierung von sequentiellen Ereignissen (Kahana, 1996), was einen solchen Abruf anspruchsvoller macht. Zweitens greifen Schemata¹⁵, die uns helfen, Ereignisse zu rekonstruieren, beim Rückwärtserzählen viel weniger (Geiselman & Callot, 1990). Für eine lügende Person ist das ein Problem, da Lügen anhand von Schemawissen konstruiert werden (siehe z.B. Volbert, 2008). Für ehrliche Personen hingegen kann das „Ausschalten“ des Einflusses von Schemawissen sogar ein Vorteil sein. Sie müssen dadurch stärker auf die Originalerinnerung zurückgreifen und erinnern so manchmal noch weitere Informationen, die sie sonst nicht aus der Erinnerung abgerufen hätten (siehe z.B. Bensi et al., 2011; Gilbert & Fisher, 2006).¹⁶ Drittens kommt die Aufforderung, das Ereignis in umgekehrter Reihenfolge zu schildern, für viele Aussagende unerwartet (z.B. Shaw et al., 2013), weshalb lügende Personen oftmals nicht darauf vorbereitet sind.

¹⁵ Volbert (2008) definiert Schemata wie folgt: „Schemata sind abstrakte Wissensstrukturen, die Vorannahmen über Gegenstände, Menschen und Situationen enthalten“ (S. 14). In einem Schema sind dabei jene Eigenschaften zusammengefasst, „die typischerweise in einem Exemplar des jeweiligen Gegenstandsbereichs vorkommen“ (ebd.).

¹⁶ Das Erzählen in einer anderen zeitlichen Abfolge ist daher auch eine Komponente des „kognitiven Interviews“ (Fisher & Geiselman, 1992), welches zur Unterstützung des Erinnerungsabrufs aussagewilliger Personen entwickelt wurde.

Tatsächlich stellten Vrij et al. (2008) beim Rückwärtserzählen mehr Unterschiede zwischen Wahrheit und Lüge fest als bei Aussagen ohne spezifische Instruktion zur Reihenfolge. Während sich im ersten Fall in acht der analysierten Hinweise statistisch signifikante Unterschiede mit mittleren bis großen Effektstärken zeigten (u.a. wurden in wahren Aussagen mehr auditive Details, mehr kontextuelle Einbettung, und weniger Blinzeln beobachtet), war im zweiten Fall nur ein Unterschied feststellbar (mehr Hand- und Armbewegungen in wahren Aussagen). Vrij et al. (2012; siehe auch Lancaster et al., 2013) konnten zudem zeigen, dass Lügen und wahre Aussagen sich stärker hinsichtlich der Menge an Details unterschieden, wenn die Ereignisse in umgekehrter Reihenfolge erzählt wurden. Berichtete dieselbe Person das Ereignis einmal chronologisch und einmal rückwärts, wurden bei lügenden Personen außerdem häufiger Diskrepanzen zwischen den beiden Aussagen festgestellt als bei ehrlichen Personen. Die Trefferquote von Beobachtern stieg in der Studie von Vrij et al. (2008) von 46% auf immerhin 58%, wenn diese rückwärts erzählte Aussagen beurteilten, wobei insbesondere Lügen besser erkannt wurden.

Zwei Punkte zur Anwendung des Rückwärtserzählens in der Praxis sind zu erwähnen. Erstens hat die Autorin die Erfahrung gemacht, dass Befragter teilweise Bedenken haben, eine Person zum Rückwärtserzählen aufzufordern, da sie die Aufforderung als unnatürlich und gegenüber dem Vernommenen schwer zu rechtfertigen empfinden. Hier kann empfohlen werden, als Begründung für die Aufforderung zum Rückwärtserzählen darauf zu verweisen, dass das Schildern der Ereignisse in einer anderen Reihenfolge beim Erinnern helfen kann, was für ehrliche Personen ja auch tatsächlich der Fall ist. Zweitens kann nicht ausgeschlossen werden, dass das Rückwärtserzählen auch eine die Wahrheit sagende Person über Gebühr fordern kann, wodurch diese dann – ähnlich einer lügenden Person – vermehrte Hinweise auf eine erhöhte kognitive Belastung zeigen würde (siehe Vrij, 2018; siehe auch 5.3.). Es müssen also jeweils Alternativerklärungen für vorliegende Hinweise überprüft werden. So könnte eine Person, die ein unterdurchschnittliches kognitives Funktionsniveau aufweist, bei einer solchen Aufforderung womöglich an ihre Grenzen stoßen, selbst wenn sie die Wahrheit sagt. Ein länger zurückliegendes Ereignis könnte ähnliche Probleme mit sich bringen, da der Erinnerungsabruf in diesem Fall anspruchsvoller sein kann.

Inhaltsbezogene unerwartete Fragen

Inhaltsbezogene unerwartete Fragen können sich beispielsweise auf zeitliche Abläufe oder räumliche Anordnungen beziehen. Bei einer Aussage zu einem Restaurantbesuch könnten unerwartete Fragen etwa lauten: „In welcher Reihenfolge haben Sie die erwähnten Themen während des Essens diskutiert?“ und „Relativ zum Eingang, wo im Lokal haben Sie und Ihre Freunde gesessen?“ (Beispiele aus Vrij et al., 2009). Das Stellen solch unerwarteter Fragen kann einerseits dann besonders nützlich sein, wenn mehrere Personen zu demselben Ereignis Auskunft geben sollen, z.B. im Rahmen einer Alibiüberprüfung, da sich die lügenden Personen zu diesen Aspekten wohl kaum vorgängig abgestimmt haben werden (Vrij et al., 2009). In der Tat fanden Vrij und Kollegen (2009), dass die Antworten zweier Personen, die den gemeinsamen Restaurantbesuch nur vorgaben, nur für unerwartete Fragen mehr Diskrepanzen aufwiesen als die Antworten zweier Personen, die tatsächlich gemeinsam Essen waren. Bei erwarteten Fragen wurde hingegen kein Unterschied beobachtet. Ferner waren bei lügenden Personen insbesondere die Antworten auf die unerwarteten

räumlichen Fragen weniger detailliert. All diese Effekte waren mittel bis groß. Andererseits können inhaltsbezogene unerwartete Fragen auch bei Aussagen von Einzelpersonen hilfreich sein. Lancaster et al. (2013) stellten fest, dass es zwar sowohl bei lügenden als auch bei ehrlichen Personen zu einer Abnahme an berichteten Details bei Antworten auf eine unerwartete im Vergleich zu einer erwarteten Frage kam, diese Abnahme bei lügenden Personen aber ausgeprägter war. Anhand dieser Reduktion in der Menge an berichteten Details bei Antworten auf unerwartete Fragen konnten mit einem statistischen Modell 78% aller wahren Aussagen und 83% aller Lügen korrekt klassifiziert werden.¹⁷

Ähnlich wie beim Rückwärtserzählen dürften auch bei inhaltsbezogenen unerwarteten Fragen die Unterschiede zwischen Lüge und Wahrheit umso geringer ausfallen, je schlechter sich ehrliche Personen an das Ereignis erinnern können. Bei der Konstruktion unerwarteter Fragen muss ferner darauf geachtet werden, dass diese von Personen, die die Wahrheit sagen, ähnlich einfach durch das Abrufen ihrer episodischen Erinnerung beantwortet werden können, wie erwartete Fragen. Daher sollten sich inhaltsbezogene unerwartete Fragen nicht auf allzu periphere Aspekte des Ereignisses beziehen.

Modalitätsbezogene unerwartete Fragen

Bei modalitätsbezogenen unerwarteten Fragen wird der Befragte zusätzlich zur (erwarteten) verbalen Auskunft auch darum gebeten, eine Skizze anzufertigen (z.B. Layout des Restaurants inkl. Informationen dazu, wo sich relevante Personen aufgehalten haben). Da so im Endeffekt dieselbe Frage zweimal beantwortet wird, kann die Übereinstimmung der beiden Antworten überprüft werden. Bei lügenden Personen wurden im Vergleich zu ehrlichen Personen deutlich mehr Inkonsistenzen zwischen der verbalen Antwort und den in der Skizze enthaltenen Informationen festgestellt (Leins et al., 2011 und 2012). Ehrlichen Personen dürfte es deshalb leichter fallen, die Ereignisse in verschiedenen Modalitäten wiederzugeben, weil sie auf ihre episodische Erinnerung an das Ereignis zurückgreifen können und dadurch kognitiv flexibler sind. Die Informationen, die eine lügende Person sich zurechtgelegt hat, sind hingegen primär auf eine verbale Wiedergabe ausgelegt (siehe z.B. Leins et al., 2012). In der Studie von Leins et al. (2011) wurden anhand der Konsistenz zwischen Aussage und Skizze mit statistischen Modellen Trefferquoten von 75% bis 89% erreicht. Bezüglich Paaren von aussagenden Personen stellten Vrij et al. (2009) fest, dass die Zeichnungen von Paaren von lügenden Personen wesentlich mehr Inkonsistenzen aufwiesen als die Zeichnungen von Paaren von ehrlichen Personen ($d = 1.10$). Statistische Modelle konnten anhand dieses Hinweises 75% der Lügen und 80% der ehrlichen Aussagen korrekt identifizieren. Weitere Forschung ist jedoch notwendig, um mehr darüber zu erfahren, ob die Validität der Hinweise beispielsweise davon abhängig ist, wie gut sich eine ehrliche Person an das Ereignis erinnern kann.

¹⁷ Trefferquoten im Sinne korrekter Klassifikationen anhand statistischer Modelle lassen nur begrenzt Rückschlüsse auf die Leistung zu, die man von einem Menschen erwarten kann. Statistische Modelle berechnen anhand der systematisch kodierten Hinweise den optimale cut-off Wert zur Unterscheidung von Lüge und Wahrheit in der vorliegenden Stichprobe und wenden diesen für die Klassifikation der Aussagen systematisch an. Das kann ein Mensch so natürlich nicht leisten.

Fazit

Die verschiedenen Arten von unerwarteten Fragen können die Unterschiede zwischen Wahrheit und Lüge vergrößern. Inhalts- und modalitätsbezogene unerwartete Fragen scheinen nicht nur bei Einzelpersonen, sondern auch im Kontext der Alibiüberprüfung hilfreich zu sein. Bei inhaltsbezogenen unerwarteten Fragen müssen die unerwarteten Fragen sorgfältig ausgewählt werden, da sie für ehrliche Personen grundsätzlich ähnlich einfach zu beantworten sein müssen wie die erwarteten Fragen. Wenn ehrliche Personen Mühe haben, sich an das Ereignis zu erinnern, z.B. weil es lange her ist oder wenig distinktiv war, könnte dies ihr Antwortverhalten und somit die Validität von Hinweisen beeinflussen. Hierzu ist weitere Forschung erforderlich.

4.3. Anregung zum detaillierteren Erzählen (Modellaussage)

Personen, die die Wahrheit sagen, berichten in einer Aussage ohne spezifische Instruktion oftmals nicht alles, was sie erinnern können (siehe z.B. Vrij et al., 2014). Deshalb, so die Idee dieses Ansatzes, sollten diese Personen im Normalfall relativ problemlos zusätzliche Details berichten können, wenn sie angeregt werden, einen detaillierteren Bericht abzugeben. Für eine zu Unrecht verdächtige Person hat eine detailliertere Aussage zudem den Vorteil, dass den Strafverfolgungsbehörden dadurch zusätzliche entlastende Informationen bekannt werden können. Demgegenüber ist das Berichten zusätzlicher Details für lügende Personen mit diversen Nachteilen behaftet (siehe z.B. Vrij, Leal & Fisher, 2018). Einerseits wollen lügende Personen glaubwürdig wirken, weshalb sie der Aufforderung des Befragers nach mehr Details nachkommen möchten.¹⁸ Andererseits ist es jedoch kognitiv anspruchsvoll, Details aus dem Stehgreif zu erfinden, was sich wiederum in der Qualität und Plausibilität der fabrizierten Information widerspiegeln könnte. Außerdem besteht die Gefahr, dass die zusätzlichen Informationen Hinweise enthalten, durch die die Lüge entlarvt werden könnte. Das Anregen zum detaillierteren Erzählen sollte also einerseits Personen, die die Wahrheit sagen, entgegenkommen und andererseits lügenden Personen ihre Aufgabe erschweren.

Der eine oder andere Leser wundert sich nun vielleicht, *weshalb* eine ehrliche Person das Geschehene in einer Vernehmung denn nicht grundsätzlich so detailliert wie möglich berichten sollte. In unserer Alltagskommunikation gehen wir davon aus, dass Gesprächsbeiträge „so informativ wie (für die gegebenen Gesprächszwecke) nötig“ und „nicht informativer als nötig“ (Grice, 1993, S. 249) ausfallen sollten (Maxime der Quantität). Der Sprecher passt den Detaillierungsgrad seiner Aussage daher oftmals der Situation an. Obwohl anzunehmen ist, dass wir bei Aussagen gegenüber den Strafverfolgungsbehörden einen höheren Detaillierungsgrad anstreben als bei einer normalen Unterhaltung, wird dennoch häufig unterschätzt, wie detailliert der Bericht tatsächlich sein sollte (Vrij et al., 2014). Strategien, die detaillierteres Erzählen anregen, helfen, die Erwartungen der aussagenden Person zu managen. Ihr wird durch diese Strategien vermittelt, dass ihre Aussage nicht nur ausgesprochen detailliert sein darf, sondern in dieser Situation auch sein soll.

¹⁸ Johnson (2006) konnte zeigen, dass reichhaltigere Aussagen von Beurteilern tatsächlich auch als glaubwürdiger wahrgenommen werden.

Personen können auf unterschiedliche Arten zu detaillierterem Erzählen angeregt werden (siehe z.B. Vrij, Fisher & Blank, 2017). Eine besonders intensiv untersuchte und daher für uns interessante Möglichkeit ist das Arbeiten mit einer sogenannten „Modellaussage“ (z.B. Boogard et al., 2014). Dabei wird der vernommenen Person eine detaillierte Aussage, welche inhaltlich nicht direkt mit dem Thema der Befragung zusammenhängt, als Modell vorgelegt (zumeist in Form einer Ton-Aufzeichnung). Die Modellaussage kann einer die Wahrheit sagenden Person helfen einzuschätzen, welcher Detaillierungsgrad in der aktuellen Situation angemessen ist. Aber auch lügende Personen verstehen natürlich die Aufforderung nach mehr Informationen und wollen dieser nachkommen, um glaubwürdig zu wirken. Es überrascht daher nicht, dass sich die reine *Menge* an zusätzlichen Details, die nach der Präsentation einer Modellaussage berichtet werden, für lügende und ehrliche Personen zumeist nicht unterscheidet (z.B. Leal et al., 2018; Vrij, Leal, Jupe & Harvey, 2018; aber siehe Porter & Salvaneli, 2020, für einen gegenteiligen Befund). Jedoch werden Unterschiede zwischen Lüge und Wahrheit ersichtlich, wenn die *Art* der berichteten Details differenzierter betrachtet wird.

Leal et al. (2018) fanden, dass lügende Personen in einem zweiten Bericht, der nach der Präsentation einer Modellaussage erfolgte, signifikant mehr zusätzliche *periphere* Details berichteten, während ehrliche Personen signifikant mehr *zentrale* Details berichteten. Dieser Befund ist nachvollziehbar, da es sowohl schwieriger als auch riskanter ist, Details zu zentralen Aspekten des Ereignisses zu erfinden. Vrij und Kollegen (Vrij, Leal et al., 2017; Vrij, Leal, Jupe & Harvey, 2018) identifizierten zudem sogenannte „Komplikationen“ als hilfreichen Hinweis. Dabei handelt es sich um für die eigentliche Haupthandlung unwichtige Details, welche die Komplexität des Ereignisses unnötig erhöhen.¹⁹ Obwohl wahre Aussagen bereits in der Kontrollbedingung ohne Modellaussage mehr Komplikationen enthielten als Lügen, war der Effekt bei Berichten, die *nach* dem Hören einer Modellaussage abgegeben wurden, wesentlich grösser ($d = 0.23$ vs. 0.75 ; Vrij, Leal et al., 2017). Ferner wurde festgestellt, dass der *Anteil*²⁰ an Komplikationen bei ehrlichen Personen in ihrem zweiten Bericht, den sie nach dem Hören einer Modellaussage abgaben, signifikant anstieg, wohingegen ein solcher Anstieg bei lügenden Personen nicht beobachtet wurde ($d = 0.80$ vs. $d = 0.08$; Vrij, Leal, Jupe & Harvey, 2018). Der Anteil an Komplikationen führte zu einer korrekten Klassifizierung von 74% der wahren und 65% der gelogenen Aussagen. Dass Komplikationen nach einer Modellaussage so nützliche Hinweise sind, kann einerseits damit

¹⁹ Komplikationen sind an das Merkmal „unerwartete Handlungskomplikationen“ aus den Realkennzeichen (Steller & Köhnken, 1989) angelehnt, sie sind jedoch breiter definiert. Eine unerwartete Handlungskomplikation im Sinne der Realkennzeichen wäre etwa, wenn man bei der Reise von A nach B die Autobahnabfahrt verpasst hat und deshalb zurückfahren musste. Die *bewusste* Entscheidung, bei einer Reise von A nach B einen Zwischenstopp in C einzulegen, würde hingegen nicht gezählt. Nach Vrij, Leal, Jupe und Harvey (2018) würde diese bewusste Entscheidung jedoch als Komplikation kodiert, da die Information für die Haupthandlung nebensächlich ist, aber gleichzeitig die Komplexität des Ereignisses erhöht.

²⁰ In dieser konkreten Studie wurde der Anteil an Komplikationen wie folgt berechnet: Komplikationen / (Komplikationen + Allgemeinwissen + selbst-benachteiligende Strategien). Allgemeinwissen bezieht sich auf stereotype Informationen zu einem Ereignis („Wir gingen ins Louvre, wo wir die Mona Lisa sahen“; Beispiel aus Vrij, Leal, Jupe & Harvey, 2018). Selbst-benachteiligende Strategien sind *Rechtfertigungen* dafür, weshalb man gewissen Informationen nicht liefern kann. Die letzten beiden Hinweise werden häufiger in Lügen als in wahren Aussagen erwartet, was in der Studie von Harvey und Kollegen (2018) auch so bestätigt werden konnte.

erklärt werden, dass sie nicht zur Strategie von lügenden Personen passen, ihre Aussage möglichst einfach zu halten (Hartwig et al., 2007). Andererseits sind Komplikationen generell schwierig zu fabrizieren, da sie nicht Teil unserer Schemata sind, anhand derer wir Lügen zumeist konstruieren. Bei einer ehrlichen Person sind Komplikationen hingegen ein normaler Bestandteil der Erinnerung an ein Ereignis. Die Modellaussage kann bei ehrlichen Personen dazu führen, dass sie vermehrt solche – für das eigentliche Hauptgeschehen nebensächliche – Informationen berichten, welche sie sonst vielleicht im Sinne der Grice'schen Maxime der Quantität als irrelevant eingestuft hätten. Schließlich soll noch auf eine Studie hingewiesen werden, welche gezeigt hat, dass wahre im Vergleich zu gelogenen Aussagen als deutlich plausibler wahrgenommen wurden, wenn den Aussagenden vor ihrem Bericht eine Modellaussage vorgelegt wurde (Leal et al., 2015). Ein statistisches Modell konnte anhand des Hinweises „Plausibilität“ 80% aller Aussagen korrekt klassifizieren. Wurde vor dem Bericht keine Modellaussage vorgelegt, unterschieden sich wahre und gelogenen Aussagen hingegen nicht in ihrer Plausibilität.

Drei praxisbezogene Punkte zur Modellaussage sollen abschließend kurz diskutiert werden. Erstens stellt sich die Frage nach den Auswirkungen einer zeitlichen Verzögerung zwischen Ereignis und Vernehmung, da in der Praxis – anders als in den meisten Studien – die Vernehmung oftmals nicht am gleichen Tag erfolgt. Wie Ebbinghaus bereits 1885 zeigen konnte, vergessen wir schon nach kurzer Zeit viele Details. Daher ist die Frage der Zeitdauer, die zwischen Ereignis und Vernehmung liegt, für die Strategie des Anregens zum detaillierteren Erzählen besonders wichtig. Was geschieht, wenn ein Ereignis lange zurückliegt und/oder der aussagenden Person dessen Wichtigkeit zum Zeitpunkt des Geschehens nicht bewusst war (z.B. Alibi für eine fälschlicherweise tatverdächtige Person)? Eine erste Studie zu den Auswirkungen einer drei-wöchigen Verzögerung zwischen Ereignis und Aussage zeigte, dass lügende Personen in einer zweiten Aussage, welche nach dem Vorlegen einer Modellaussage erfolgte, im Vergleich zu ihrer ersten Aussage signifikant *mehr* zusätzliche Details berichten als ehrliche Personen ($d = 0.61$) (Harvey et al., 2019). Die Autoren erklären dies damit, dass bei ehrlichen Personen nach drei Wochen normale Vergessensprozesse ihren Tribut forderten. Dadurch, dass generell weniger Details erinnert wurden, wurde die meiste Information bereits im ersten Bericht wiedergegeben. Lügende Personen hingegen versäumten es, bei ihrer Aussage diesen natürlichen Prozess zu berücksichtigen und zeigten einen *Stabilitäts-Bias* (Harvey et al., 2019). Viele zusätzliche Details in einem zweiten Bericht, der nach dem Vorlegen einer Modellaussage erfolgt, können also ein Hinweis auf eine *Lüge* sein, falls aus irgendwelchen Gründen davon ausgegangen werden muss, dass die Erinnerung von ehrlichen Personen an das Ereignis eher schwach ausfallen könnte. Zweitens sollte beim Einsatz dieser Strategie berücksichtigt werden, dass sich Personen darin unterscheiden, wie detailliert ihre Erzählungen normalerweise sind. So kann zumindest nicht ausgeschlossen werden, dass einige ehrliche Personen in einem ersten freien Bericht die Ereignisse bereits so ausführlich schildern, dass eine Steigerung kaum noch möglich ist, wenn sie gebeten werden, nach dem Vorlegen einer Modellaussage die Geschehnisse nochmals zu berichten. Drittens sollte gut überlegt sein, wie das Einholen zweier Berichte (einer vor und einer nach der Modellaussage) sowie das Vorlegen der Modellaussage selbst sinnvoll erklärt werden können. Der Vernommene könnte beispielsweise dahingehend informiert werden, dass es Personen oft schwerfällt abzuschätzen, wie detailliert ihre Aussage sein soll, und dass die Modellaussage sie bei ihrer

Einschätzung unterstützen und ihnen das Aussagen somit erleichtern kann (siehe z.B. Harvey et al., 2019).

Fazit

Insgesamt ist das Anregen detaillierteren Erzählens unter Zuhilfenahme einer Modellaussage eine vielversprechende Strategie. Sie kann nicht nur Unterschiede zwischen Wahrheit und Lüge verstärken, sondern unter Umständen aufgrund der ausführlicheren Aussagen ehrlicher Personen auch zusätzliche für die Ermittlungen relevante Informationen ans Licht bringen. Befrager sollten allerdings Vorsicht walten lassen, wenn das Ereignis schon etwas länger zurückliegt oder zum Zeitpunkt des Geschehens von der aussagenden Person nicht als relevant erkannt wurde. Dann ist es möglich, dass lügende Personen sogar mehr zusätzliche Informationen berichten als ehrliche, was die Gefahr falsch-positiver Urteile (eine ehrliche Person wird fälschlicherweise als Lügner klassifiziert) mit sich bringt.

4.4. Nachprüfbarkeits-Ansatz

Wie unter 4.3. bereits erörtert, versucht eine lügende Person oftmals, ihre Aussage mit Details anzureichern, da sie sich der Tatsache bewusst ist, dass eine detailliertere Erzählung als glaubwürdiger wahrgenommen wird (Hartwig et al., 2007; Nahari et al., 2012). Gleichzeitig will sie aber auch vermeiden, dass sie aufgrund der zusätzlichen Details der Lüge (z.B. falsches Alibi) überführt werden könnte. Die zentrale Annahme des Nachprüfbarkeits-Ansatzes („verifiability approach“; Nahari et al., 2014a) ist, dass lügende Personen dieses Dilemma lösen, indem sie gezielt Details schildern, die ihrer Meinung nach *nicht* nachprüfbar sind. Dadurch können sie detailliertere Aussagen generieren, ohne dabei die Gefahr zu erhöhen, dass ihre Lügen bei einer Überprüfung offenbart werden könnten. Bei Personen, die die Wahrheit sagen, spielen solche Erwägungen hingegen keine Rolle. Vielmehr ist es sogar in ihrem Interesse, dass der Befrager erfährt, was wirklich vorgefallen ist. Personen, die die Wahrheit sagen, sind also motivierter, verifizierbare Details zu berichten, und sie sind auch eher dazu in der Lage (siehe Vrij & Nahari, 2018, S. 128).²¹

Es wird daher erwartet, dass ehrliche Personen im Vergleich zu lügenden Personen im Schnitt mehr nachprüfbare Details berichten und dass ihre Aussagen auch einen höheren *Anteil*²² an nachprüfbaren Details aufweisen als die Aussagen von lügenden Personen (Nahari et al., 2014b; Vrij & Nahari, 2018). Als nachprüfbar gelten alle Details, die (a) mit einer identifizierbaren Person erlebt, (b) von einer identifizierbaren Person beobachtet oder (c) durch Technologie aufgezeichnet oder anderweitig dokumentiert wurden (z.B. Überwachungsvideo, Telefongespräch, Anwesenheitsliste; Nahari, 2019). Die Analyse der Aussagen hinsichtlich des Vorliegens von verifizierbaren und nicht verifizierbaren Details

²¹ Dies setzt natürlich voraus, dass das Ereignis grundsätzlich verifizierbare Aspekte enthält. Mit den technologischen Entwicklungen gibt es mittlerweile für erstaunlich viele Situationen potentiell überprüfbare Details. So kann beispielsweise die Aussage, man sei alleine zu Hause gewesen und früh zu Bett gegangen, theoretisch überprüfbare Informationen zur Nutzung von Mobiltelefon, Computer oder Fernseher enthalten (Beispiel aus Vrij & Nahari, 2018, S. 129f).

²² Gemessen am Total der berichteten nachprüfbaren und nicht nachprüfbaren Details.

erfolgt durch einen geschulten Kodierer.²³ Obwohl die Definition nachprüfbarer Details an sich gut nachvollziehbar sein dürfte, gilt es zu berücksichtigen, dass die potentielle Nachprüfbarkeit eines Details nur dann diagnostisch ist, wenn eine lügende Person sich der theoretischen Überprüfbarkeit einer Information auch bewusst ist und diese Information somit als zu vermeidend einstufen würde. Diesem Umstand wird Rechnung getragen, indem mit einer „reasonable standard person assumption“ gearbeitet wird, d.h., es wird beurteilt, ob eine vernünftige Durchschnittsperson eine Information als verifizierbar erachten würde oder nicht (Nahari, 2019).

Wie ebenfalls bereits unter 4.3. diskutiert, haben ehrliche Personen oftmals Schwierigkeiten damit einzuschätzen, wie detailliert ihre Aussage sein soll, und berichten daher nicht alles, was sie erinnern können. Aus diesem Grund werden Aussagende im Rahmen des Nachprüfbarkeits-Ansatzes instruiert, möglichst *alle* Details zu schildern, auch wenn sie irrelevant erscheinen (Nahari, 2019). Empirische Befunde zum Einsatz des Nachprüfbarkeits-Ansatzes in einem polizeilich relevanten Szenario²⁴ haben gezeigt, dass Personen, die die Wahrheit sagen, in der Tat mehr verifizierbare Details berichteten als lügende Personen (z.B. Nahari et al., 2014a; Vernham et al., 2020; aber siehe z.B. Jupe et al., 2020). Interessanterweise wurde zudem festgestellt, dass diese Unterschiede noch verstärkt werden konnten, wenn eine erweiterte Instruktion verwendet wurde, bei der die Person zusätzlich zur „Alles-Erzählen-Instruktion“ auch explizit über den Nachprüfbarkeits-Ansatz in Kenntnis gesetzt und darum gebeten wird, möglichst viele *nachprüfbare* Details zu schildern (Jupe et al., 2020; Nahari et al., 2014b; siehe auch Vrij & Nahari, 2018). Ein naheliegender Grund für diesen Befund ist, dass ehrliche Personen durch die erweiterte Instruktion gezielter darüber nachdenken, ob etwas Nachprüfbares vorgefallen ist, wohingegen lügende Personen nicht von dieser Anregung profitieren können. Nach aktuellem Stand der Forschung wird daher empfohlen, die Aussagenden zusätzlich zu einer „Alles-Erzählen-Instruktion“ dazu aufzufordern, möglichst viele Details zu berichten, die ihrer Meinung nach vom Befrager überprüft werden können (Vrij & Nahari, 2018).

Über Studien mit der einfachen und der erweiterten Instruktion hinweg berichten Vrij und Nahari (2018) für den Hinweis „Anzahl nachprüfbarer Details“ große Effekte, mit Cohens *d* zwischen 0.70 und 2.05.²⁵ Auch mit dem *Anteil* an verifizierbaren Details konnten in zwei von drei Studien gelogene von wahren Aussagen gut unterschieden werden. Mit statistischen Modellen konnten anhand der Anzahl nachprüfbarer Details 59% bis 88% der Aussagen korrekt klassifiziert werden (Vrij & Nahari, 2018). Der Anteil nachprüfbarer Details führte ebenfalls zu guten Klassifikationsraten von 67% und 85%. Untrainierte Laien hingegen, die dazu angehalten wurden, auf verifizierbare Details zu achten und diese in

²³ Es wird auch an einer Automatisierung der Kodierung gearbeitet, jedoch steht noch keine einsatzbereite Software o.ä. zur Verfügung (Kleinberg et al., 2016, zitiert in Nahari, 2019).

²⁴ Studien, die den Nutzen des Nachprüfbarkeits-Ansatzes bei Versicherungsmeldungen oder bei Simulantentum untersuchten, werden hier nicht besprochen. Für eine Übersicht über die Ergebnisse dieser Studien, siehe Vrij und Nahari (2018).

²⁵ Die meisten der bisher durchgeführten Studien stammen von derselben Forschergruppe. Eine andere Forschergruppe konnte in ihrer Studie die Befunde zur Validität der Anzahl verifizierbarer Details nicht bestätigen (Bogaard et al., 2020). Allerdings sind die Studien aufgrund unterschiedlicher Herangehensweisen nicht ohne Weiteres miteinander vergleichbar, und die genaue Ursache der fehlenden Replikation bleibt unklar. Weitere Studien von unabhängigen Forschungsgruppen zu diesem Ansatz wären aber auf alle Fälle wünschenswert.

ihrem Urteil zu berücksichtigen, erreichten lediglich eine Trefferquote von 56%, wobei der Wert je nach Bedingung stark variierte (Vernham et al., 2020). Das aufwendige Kodieren scheint nach diesem ersten Befund also nicht durch die einfache Instruktion, auf verifizierbare Details zu achten, ersetzt werden zu können.

Für einen Einsatz der Strategie in der Praxis ist eine Überprüfung des Wahrheitsgehalts der nachprüfbarer Details theoretisch gesehen nicht erforderlich. Der Grund dafür liegt in der Annahme, dass lügende Personen gar nicht erst nachprüfbare Details fabrizieren, aus Angst, ihre Aussage würde einer Überprüfung nicht standhalten (Vrij & Nahari, 2018). Das Vorliegen vieler nachprüfbarer Details legt daher nahe, dass es sich um eine wahre Aussage handelt. Dennoch kann natürlich nicht ausgeschlossen werden, dass es lügende Personen gibt, die gerade darauf bauen, dass ihre Angaben nicht näher überprüft werden. Es gilt im Einzelfall also immer sorgfältig abzuwägen, ob ein Überprüfen der Angaben nicht doch angezeigt ist, sollte eine solche Verifizierung nicht ohnehin standardmäßig stattfinden (siehe auch Vrij & Nahari, 2018). Bei der Entwicklung neuer Befragungsstrategien stellt sich zudem auch immer die Frage der Übertragbarkeit der Ergebnisse vom Labor auf die Praxis. In Falle des Nachprüfbarkeits-Ansatzes ist es interessanterweise möglich, dass der Effekt in der Praxis sogar *stärker* ausfallen könnten als in Laborstudien, da lügende Personen umso weniger (fabrizierte) potentiell nachprüfbare Details berichten sollten, je höher sie die Wahrscheinlichkeit einer tatsächlichen Überprüfung einschätzen (siehe auch Vrij & Nahari, 2018). Eine Überprüfung dieser Annahme steht allerdings noch aus. Schließlich soll auch hier kurz auf die Frage nach den Auswirkungen einer zeitlichen Verzögerung zwischen Ereignis und Befragung eingegangen werden. Die Befunde einer ersten Studie mit einer 7-tägigen Verzögerung zwischen Ereignis und Aussage legen nahe, dass die Anzahl nachprüfbarer Details in einem solchen Fall möglicherweise weniger gut zwischen Wahrheit und Lüge diskriminieren kann (Jupe et al., 2020). So war lediglich eine ganz spezifische Art von verifizierbaren Details valide (ehrliche Personen berichteten signifikant mehr zusätzliche nachprüfbare digitale Quellen-Details), während für das Total an verifizierbaren Details keine signifikanten Unterschiede beobachtet wurden. Es scheint also, dass auch beim Nachprüfbarkeits-Ansatz die Stärke der Erinnerung einer ehrlichen Person ein wichtiger Faktor ist, der berücksichtigt werden muss.

Fazit

Insgesamt handelt es sich beim Nachprüfbarkeits-Ansatz um ein vielversprechendes Verfahren. Mit einer einfachen zusätzlichen Instruktion im Rahmen des informationsammelnden Befragungsstils kann der dadurch gewonnene freie Bericht anhand des Ansatzes analysiert werden. Ein Vorteil ist, dass – insbesondere bei der Verwendung der erweiterten Instruktion – allenfalls auch zusätzliche Beweise oder Indizien (im Sinne nachprüfbarer Informationen) generiert werden, die im Rahmen der Ermittlungen relevant sein können (siehe auch Vrij & Nahari, 2018). Allerdings ist die manuelle Kodierung zeitaufwendig und erfordert Training, und es herrscht auch ein gewisses Maß an Subjektivität, da die Kategorisierung als „verifizierbar“ die Annahme voraussetzt, dass ein „vernünftiger Durchschnittsmensch“ diese Information als nachprüfbar einschätzen würde. Schließlich ist bei der Anwendung des Ansatzes Vorsicht geboten, wenn es Hinweise darauf gibt, dass sich eine ehrliche Person möglicherweise nicht (mehr) so gut an das Ereignis erinnern kann (z.B. das Ereignis liegt länger zurück oder wurde zum damaligen Zeitpunkt nicht als wichtig erkannt).

4.5. Strategische Nutzung von Evidenz

Die Technik der strategischen Nutzung von Evidenz („strategic use of evidence technique“; kurz SUE-Technik; Hartwig et al., 2005) kann dann eingesetzt werden, wenn dem Befrager belastende Indizien oder Beweise gegen die zu vernehmende Person vorliegen. Die SUE-Technik basiert auf der Annahme, dass sich zu Recht und zu Unrecht Verdächtige in den Strategien unterscheiden, die sie während der Befragung einsetzen, um den Befrager von ihrer Unschuld zu überzeugen (Hartwig et al., 2005). Konkret wird erwartet, dass sich schuldige und unschuldige Personen²⁶ in ihrem Informationsmanagement unterscheiden (siehe z.B. Granhag & Hartwig, 2015), also darin, welche Informationen sie dem Befrager offenbaren und welche nicht. Schuldige Personen haben Tatwissen, welches sie vor dem Befrager verbergen müssen. Dies können sie entweder durch eine Vermeidungsstrategie (z.B. Anwesenheit am Tatort wird nicht erwähnt) oder durch eine Flucht- bzw. Verneinungsstrategie (z.B. Anwesenheit am Tatort wird verneint) erreichen. Im Gegensatz dazu liegt es gerade im Interesse unschuldiger Personen, dass der Befrager erfährt, was tatsächlich vorgefallen ist. Daher wird erwartet, dass unschuldige Personen die Ereignisse wahrheitsgemäß schildern sowie kooperativ und zuvorkommend sind.

Die Annahme, dass unschuldige Personen die Ereignisse wahrheitsgemäß schildern – und zwar selbst dann, wenn sie sich dadurch potentiell selbst belasten, weil sie beispielsweise tatsächlich am Tatort waren, aber aus einem anderen Grund – basiert auf zwei psychologischen Phänomenen: der Illusion der Transparenz (Gilovich et al., 1998) und dem Gerechte-Welt-Glauben (Lerner, 1980). Die Illusion der Transparenz beschreibt die Beobachtung, dass wir oftmals überschätzen, wie wahrnehmbar unsere inneren Vorgänge und Emotionen für andere Personen sind. Für unschuldige Tatverdächtige bedeutet dies,

²⁶ Die Begriffe „schuldige“ und „unschuldig“ sind im umgangssprachlichen Sinn zu verstehen, d.h., ob eine Person die Tat, derer sie verdächtigt wird, tatsächlich begangen hat oder nicht. Ferner sei darauf hingewiesen, dass im Kontext dieses Ansatzes aufgrund der theoretischen Annahmen „Unschuld“ mit „Wahrheit“ und „Schuld“ mit „Lüge“ mehr oder weniger gleichgesetzt werden.

dass sie das Gefühl oder sogar die Überzeugung haben, man müsse ihnen ihre Unschuld letztlich „ansehen“ (Granhag & Hartwig, 2015). Der Gerechte-Welt-Glaube beschreibt, wie der Name bereits sagt, unseren Glauben daran, dass die Welt ein gerechter Ort ist, an dem eine Person das bekommt, was sie verdient. Ich als unschuldige, zu Unrecht einer Straftat verdächtige Person, die kooperativ ist und der Polizei alle Informationen zur Verfügung stellt, muss am Ende also entlastet werden.²⁷

Aufgrund dieser theoretischen Überlegungen wird nun angenommen, dass sich die Aussage einer schuldigen Person von jener einer unschuldigen Person systematisch unterscheidet, und zwar sowohl im freien Bericht als auch in ihren Antworten auf beweismittelbezogene Fragen. Beweismittelbezogene Fragen sind spezifische Fragen, welche vor dem Hintergrund der vorhandenen Beweise gestellt werden, ohne diese Beweise offenzulegen (vgl. z.B. Hartwig et al., 2007). Konkret werden die folgenden zwei Unterschiede erwartet: (1) Unschuldige Personen bringen in einem freien Bericht eher potentiell selbstbelastende Informationen freiwillig (d.h. ohne spezifische Nachfrage) zur Sprache als schuldige Personen, von denen erwartet wird, dass sie eine Vermeidungsstrategie verfolgen, und (2) schuldige Personen widersprechen beim Beantworten beweismittelbezogener Fragen den vorhandenen Beweisen aufgrund ihrer Verneinungsstrategie eher als unschuldige Personen, die sich insgesamt zuvorkommend (im Sinne von offen/ehrlich) verhalten sollten. Daher sollte die „Aussage-Evidenz-Inkonsistenz“ bei schuldigen Personen höher ausfallen.

Diese Unterschiede werden nun insbesondere dann erwartet, wenn der aussagenden Person die vorhandenen Beweise zu Beginn der Vernehmung nicht bekannt sind. Denn weiß eine schuldige Person beispielsweise, dass die Polizei Kenntnis ihrer Anwesenheit am Tatort hat, könnte sie dies in ihrer Aussage berücksichtigen (z.B. Hartwig et al., 2005). So könnte sie beispielsweise zugeben, am Tatort gewesen zu sein, dafür aber eine alternative Erklärung liefern. Dadurch würden sich die Antworten von schuldigen und unschuldigen Tatverdächtigten ähnlicher. Bei einer SUE-Befragung legt man die bekannten Indizien oder Beweise nicht am Anfang der Vernehmung offen, sondern erst gegen Schluss, wenn es darum geht, allfällige Widersprüche zwischen der Aussage und der Polizei anderweitig bekannten Informationen zu klären. Mit diesem Vorgehen kann der diagnostische Wert von Beweisen maximiert werden.

Die SUE-Befragung sieht zunächst einen freien Bericht vor, bei dem der Befragte gebeten wird, seine Version der Dinge zu schildern. Die Antworten auf die offenen Fragen können im Hinblick auf das freiwillige Vorbringen von potentiell selbstbelastenden Informationen analysiert werden. In einem nächsten Schritt werden beweismittelbezogene Fragen gestellt. Weiß man beispielsweise, dass eine Person an einem bestimmten Ort war, den sie noch nicht erwähnt hat, bieten sich Fragen an wie „Haben Sie alle Orte erwähnt, an denen Sie sich im relevanten Zeitraum aufgehalten haben?“, „Kennen Sie den Ort X?“, „Waren Sie zur fraglichen Zeit in der Nähe von Ort X?“. Anhand solcher beweismittelbezogener Fragen soll in einem ersten Schritt überprüft werden, ob die Aussage der Person mit der vorhandenen

²⁷ Obwohl die Annahmen zur Rolle der „Illusion der Transparenz“ und dem „Gerechte-Welt-Glauben“ mit den von unschuldigen Verdächtigten berichteten Motivatoren für ihr Verhalten übereinstimmen (z.B. Hartwig et al., 2007; Kassin, 2005), lieferte eine Studie, welche den Einfluss des Gerechte-Welt-Glaubens direkt zu testen versuchte, keine klaren Ergebnisse (Jordan & Hartwig, 2013). Es bedarf also weiterer Forschung zur Frage, was unschuldige Personen dazu motiviert, freiwillig Informationen offenzulegen, die sie selbst belasten könnten.

Evidenz übereinstimmt. Wenn nicht, geht es in einem zweiten Schritt darum, anhand der gestellten Fragen mögliche alternative Erklärungen auszuschließen, die der Befragte später als Grund für die Diskrepanz zwischen seiner Aussage und den vorhandenen Beweisen heranziehen könnte. Als Beispiel: Aufgrund eines Radarfotos ist bekannt, dass das Auto einer Person zum Zeitpunkt Y am Ort X war. Der Fahrer des Wagens ist auf dem Foto jedoch nicht klar erkennbar. Hier könnte die Person behaupten, den Wagen an diesem Tag nicht selbst gefahren, sondern ihn jemandem geliehen zu haben. Hat man diese Möglichkeit nun vorgängig bereits erfragt, wäre das Vorbringen dieser alternativen Erklärung nicht mehr besonders glaubhaft.

Diverse Studien haben gezeigt, dass sich die SUE-Technik im Vergleich zu einer Standard-Vernehmung mit früher Offenlegung von Beweisen in der Tat positiv auf den diagnostischen Wert der beiden Hinweise „freiwilliges Vorbringen von potentiell selbstbelastenden Informationen“ und „Aussage-Evidenz-Inkonsistenz“ auswirkt. In ihrer Meta-Analyse berichten Hartwig et al. (2014) unter Einbezug von acht „mock crime“-Studien²⁸, dass die Effektstärke der Aussage-Evidenz-Inkonsistenz von $d = 1.06$ in der Kontrollbedingung mit früher Offenlegung der Evidenz auf $d = 1.89$ stieg, wenn die SUE-Technik genutzt wurde. Auch hinsichtlich des freiwilligen Vorbringens von potentiell selbstbelastenden Informationen finden sich oftmals Unterschiede (Hartwig et al., 2005 und 2006, aber siehe Jordan et al., 2012), jedoch sind sie insgesamt etwas schwächer ausgeprägt (Hartwig et al., 2011). Sowohl trainierte Befrager als auch untrainierte Beobachter scheinen von den durch die SUE-Technik hervorgebrachten Hinweisen profitieren zu können. Hartwig et al. (2005) berichteten, dass die Trefferquote von untrainierten Beobachtern von 42% auf 61% anstieg, wenn statt einer Standard-Befragung die SUE-Technik verwendet wurde. Für trainierte Beobachter stellten Hartwig et al. (2006) sogar einen Anstieg von 56% auf 85% fest, wobei der Effekt in einer vergleichbaren Studie weniger ausgeprägt war (von 43% auf 65%; Luke, Hartwig, Joseph et al., 2016). Schließlich scheint die SUE-Technik gemäß ersten Befunden für Gegenmaßnahmen durch schuldige Personen nicht besonders anfällig zu sein (Luke, Hartwig, Shamash & Granhag, 2016). Obwohl über die Funktionsweise der SUE-Technik informierte schuldige Personen etwas zuvorkommender waren als nicht informierte schuldige Personen, waren sie immer noch wesentlich weniger zuvorkommend als unschuldige Personen.

In den ersten Studien zur SUE-Technik wurde kaum beobachtet, dass Unschuldige der Evidenz widersprochen hätten. Mittlerweile gibt es jedoch mindestens zwei Studien, in denen ein nicht unerheblicher Anteil der unschuldigen Personen – in einem Fall immerhin ein Drittel – sich nicht entsprechend den theoretischen Annahmen verhalten und der vorhandenen Evidenz zumindest teilweise widersprochen hat (Luke, Hartwig, Joseph et al., 2016; siehe auch Jordan et al., 2012). Die Gründe dafür konnten bisher nicht abschließend geklärt werden. Damit die Anwendung der SUE-Technik nicht zu Fehleinschätzungen durch den Befrager führt, ist es daher von größter Wichtigkeit, dass immer auch alternative Erklärungen für eine hohe Aussage-Evidenz-Inkonsistenz in Betracht

²⁸ Bei einer „mock crime“-Studie werden die Teilnehmenden instruiert, entweder ein Delikt zu begehen (z.B. etwas aus einer Aktentasche stehlen, die unbeaufsichtigt in einer Videothek steht) oder aber einer unschuldigen alternativen Aktivität nachzugehen (z.B. sich in einer Videothek nach einer bestimmten DVD umsehen). Danach sollen alle Teilnehmer den Befrager davon überzeugen, dass Sie das Delikt nicht begangen haben.

gezogen werden. Aufgrund der theoretischen Annahmen zum Verhalten unschuldiger Verdächtigter stellt sich zudem die Frage, was geschieht, wenn eine unschuldige Person *nicht* davon überzeugt ist, dass der Befrager selbst dann an ihre Unschuld glauben wird, wenn sie potentiell selbst-belastende Informationen preisgibt. Dies könnte insbesondere unschuldige Tatverdächtige aus ethnischen Minderheiten betreffen, welche zudem teilweise auch einen weniger ausgeprägten Gerechte-Welt-Glauben zu haben scheinen (für eine Diskussion dieser Möglichkeit, siehe z.B. Jordan et al., 2012). Diese Frage könnte in der Praxis noch viel schwerer wiegen, denn im Vergleich zu der Situation in wissenschaftlichen Studien steht hier für unschuldige Tatverdächtige viel mehr auf dem Spiel, wenn es ihnen nicht gelingen sollte, den Befrager von ihrer Unschuld zu überzeugen.

Fazit

Die SUE-Technik setzt voraus, dass Indizien oder Beweise vorliegen, welche nicht vorgängig offengelegt werden mussten und die der befragten Personen auch nicht anderweitig bekannt geworden sind (z.B. aus den Medien). Die Technik ist im Rahmen eines informationssammelnden Befragungsstils relativ einfach umsetzbar; lediglich die beweismittelbezogenen Fragen müssen vorbereitet werden. Liegen Widersprüche zwischen den Beweisen und der Aussage vor, ist es aber zentral, dass der Ermittler immer alternative Erklärungsmöglichkeiten berücksichtigt, da manchmal auch unschuldige Personen nicht die (ganze) Wahrheit zu sagen scheinen.

5. Möglichkeiten und Fallstricke bei der Anwendung spezifischer Befragungsstrategien in der Praxis

Zum Schluss wollen wir uns einigen Fragen zuwenden, welche sich bei einer möglichen Anwendung der spezifischen Befragungsstrategien in der Praxis stellen. Wie in Kapitel 2 dargelegt, ist ein informationssammelnder Befragungsstil im Sinne einer übergeordneten Herangehensweise zu empfehlen, wobei spezifische Befragungsstrategien im Rahmen des informationssammelnden Befragungsstils eingesetzt werden können. Konkret wollen wir uns nun überlegen, was bei der Anwendung dieser Strategien in der Praxis berücksichtigt werden muss, welches Potential die Ansätze haben und welche Fallstricke vermieden werden sollten.

5.1. Was können spezifische Befragungsstrategien gemäß Studien grundsätzlich leisten?

Einige der besprochenen Befragungsstrategien führten bei untrainierten, zumeist neutralen Beobachtern zu einer Verbesserung der Trefferquote (z.B. Rückwärtserzählen, Anregung zum detaillierteren Erzählen, SUE-Technik). Diese Befunde deuten darauf hin, dass Beobachter die durch die Befragungsstrategien hervorgerufenen Hinweise zumindest bis zu einem gewissen Grad intuitiv, also ohne spezifisches Training, nutzen können. Somit scheint bereits der bloße Einsatz einer Strategie zu einer Verbesserung der Einschätzung des Wahrheitsgehalts der Aussagen führen zu können, ohne dass eine eingehende Analyse der verbalen Hinweise durch den Vernehmenden notwendig wäre. Die erreichten Trefferquoten liegen jedoch teilweise nach wie vor unter 70%, was im Umkehrschluss eine nicht

unerhebliche Fehlerquote von mindestens 30% bedeutet (siehe z.B. Vrij, Fisher & Blank, 2017). Eine Ausnahme bildet der Nachprüfbarkeits-Ansatz, bei dem nach ersten Ergebnissen eine erfolgreiche Nutzung der Hinweise sogar *ausschließlich* mit einer systematischen Analyse möglich ist, da Beobachter selbst mit spezifischen Instruktionen keine besonders hohen Trefferquoten erzielten (Vernham et al., 2020).

Studienergebnisse zeigen aber auch, dass Beurteiler intuitiv nicht das volle Potential der vorhandenen Hinweise nutzen können. So waren die Trefferquoten oftmals deutlich höher, wenn im Vergleich die wahren und gelogenen Aussagen anhand statistischer Modelle, welche die zuvor kodierten Hinweise nutzen, klassifiziert wurden (obwohl die Trefferquoten auch hier deutlich unter 100% liegen). Wenn also in der Praxis zusätzlich eine systematische Analyse der in einer Aussage vorliegenden Hinweise durchgeführt würde, so könnte dies das Potential der Anwendung von spezifischen Befragungsstrategien wahrscheinlich noch erhöhen. Dazu wäre allerdings eine Kodierung der Aussagen erforderlich, was je nach zu kodierendem Hinweis mehr oder weniger komplex und aufwändig ist. Während es relativ einfach sein dürfte, eine Aussage systematisch auf ihre Übereinstimmung mit den vorliegenden Indizien und Beweisen zu überprüfen, ist die Kodierung von nachprüfbar und nicht nachprüfbar Details eine komplexere Aufgabe. Für eine systematische Kodierung ist zudem eine Dokumentation der Vernehmung erforderlich, welche wortgetreu und lückenlos alle Fragen und Antworten festhält. Aber selbst bei einer optimalen Kodierung werden die Trefferquoten, die man anhand statistischer Modelle findet, in der Praxis wohl nicht erreicht werden können. Bei statistischen Modellen wird nämlich *anhand der vorliegenden Fälle* der optimale cut-off Wert (im Sinne einer Entscheidungsregel) für die vorliegenden Fälle berechnet. Ein solcher cut-off Wert liegt in der Praxis jedoch nicht vor (siehe auch 5.2.). Nichtsdestotrotz ist anzunehmen, dass eine systematischere Analyse zu insgesamt korrekteren Urteilen führen dürfte als eine rein „intuitive“ Nutzung, wenn auch in etwas geringerem Ausmaß.

Fazit

Während für einige Befragungsstrategien bereits eine gewisse Verbesserung der Glaubhaftigkeitseinschätzung durch deren bloßen Einsatz möglich ist („intuitive“ Nutzung der Hinweise), braucht es bei anderen – insbesondere dem Nachprüfbarkeits-Ansatz – nach aktuellem Kenntnisstand eine systematischere Analyse der Aussagen. Für alle Strategien ist es aber wahrscheinlich, dass eine systematische Auswertung der Hinweise zu korrekteren Einschätzungen führt als eine intuitive Nutzung. Die Zuverlässigkeit einer Einschätzung des Wahrheitsgehalts einer Aussage darf jedoch so oder so nicht überschätzt werden.

5.2. Übertragbarkeit von Studienbefunden in die Praxis

Eine wichtige Frage ist, inwieweit die Ergebnisse aus den Studien nun konkret auf den Einzelfall in der Praxis angewendet werden können. In Studien erfolgen Vergleiche auf Gruppenebene. D.h., eine Gruppe von lügenden Personen wird mit einer Gruppe von Personen verglichen, die die Wahrheit sagen. Dabei sind das Ereignis, über das die Vernehmenden eine Aussage machen, die Dauer, die seit dem Ereignis vergangen ist, sowie

die Art der Befragung standardisiert, so dass alle Aussagenden relativ ähnliche Bedingungen antreffen. Die Ergebnisse der Studien geben also Auskunft darüber, wie sich lügende Personen *unter diesen spezifischen Bedingungen im Durchschnitt* von Personen unterscheiden, welche die Wahrheit sagen. In der Praxis ist die Situation hingegen eine gänzlich andere. Hier haben wir einen sogenannten „Einzelfall“, bei dem es gilt, den Wahrheitsgehalt der Aussage einer einzelnen Personen einzuschätzen.

Im Folgenden stellen wir zunächst Überlegungen dazu an, wie sich die Situation in Laborstudien von jenen in echten Vernehmungen unterscheidet, und welche Auswirkungen diese Unterschiede für die Übertragbarkeit der Ergebnisse in die Praxis haben könnten. Sodann wenden wir uns den Schwierigkeiten zu, die mit einer Einzelfall-Beurteilung verbunden sind.

Unterschiede zwischen Laborstudien und echten Vernehmungen

Bei der Frage, wie gut Laborstudien die Verhältnisse einer echten Vernehmung widerspiegeln, sind zwei Punkte zu beachten:

Erstens dürfte für die aussagende Person in der Laborsituation im Allgemeinen weniger auf dem Spiel stehen, als wenn sie sich in einer echten Vernehmung befindet. Dies gilt sowohl für die lügende Person für den Fall, dass ihre Lüge auffliegt, als auch für die ehrliche Person, die mit dem Risiko konfrontiert ist, dass man ihr nicht glaubt. Aus ethischen Gründen können Studienteilnehmenden, die den Befrager nicht von der Wahrheit ihrer Aussage überzeugen können, keine vergleichbar gravierende Konsequenzen angedroht werden, wie sie in der Realität vorliegen. Dieser Unterschied ist nach Meinung der Autorin insbesondere für die SUE-Technik von Relevanz. Die für eine unschuldige Person entstehenden Konsequenzen, wenn ihr in einer echten Vernehmung fälschlicherweise nicht geglaubt wird, könnten bei einigen unschuldigen Personen ein ähnliches Antwortverhalten hervorrufen, wie es sonst bei schuldigen Personen beobachtet wird. Leider wissen wir bisher wenig darüber, welche Faktoren die Wahrscheinlichkeit eines solchen Verhaltens bei unschuldigen Personen erhöhen. Aber auch in Hinblick auf Befragungsstrategien, bei denen das Berichten bestimmter Details von Bedeutung ist (Nachprüfbarkeits-Ansatz, Anregen detaillierteren Erzählens, unerwartete Fragen), könnten die in der Praxis schwerwiegenderen möglichen Konsequenzen problematische Auswirkungen haben. Konkret könnte sich der durch die Vernehmung ausgelöste Stress negativ auf das Erinnerungsvermögen von ehrlichen Personen auswirken (siehe z.B. Shields et al., 2017), was den Erfolg der genannten Befragungsstrategien beeinträchtigen würde. Umso wichtiger ist es, dass die Grundsätze des informationssammelnden Befragungsstils in einer Vernehmung befolgt werden (z.B. Unvoreingenommenheit, Beziehungsaufbau, aktives Zuhören), so dass eine zu Unrecht verdächtige Person möglichst gute Bedingungen für einen erfolgreichen Erinnerungsabruf vorfindet.

Zweitens ist die Zeitdauer, die zwischen Ereignis und Vernehmung liegt, bei Laborstudien oftmals kürzer als in der Realität. Bei Laborstudien findet die Befragung nicht selten unmittelbar nach dem Ereignis statt. Außerdem ist den Teilnehmenden zumeist bereits dann, wenn sie das Ereignis erleben, bekannt, dass sie später dazu befragt werden. Dies kann dazu führen, dass dem Ereignis mehr Aufmerksamkeit geschenkt wird, als es

womöglich sonst der Fall wäre, wodurch dieses später auch besser erinnert werden kann (siehe z.B. Unsworth & Spillers, 2010). Wie bereits angesprochen, sind die Fragen der zeitlichen Verzögerung und des Bewusstseins der Bedeutsamkeit eines Ereignisses zum Wahrnehmungszeitpunkt insbesondere für jene Befragungsstrategien von großer Bedeutung, bei denen die *Detailliertheit* der Erinnerung in der einen oder anderen Form als Hinweis eine Rolle spielt. Dazu gehören die Strategien des Stellens von unerwarteten Fragen und der Anregung zum detaillierteren Erzählen sowie der Nachprüfbarkeits-Ansatz. Erfreulicherweise haben mittlerweile erste Studien untersucht, wie sich eine zeitliche Verzögerung und der Grad der Aufmerksamkeit zum Wahrnehmungszeitpunkt auf den Detailgehalt einer Erzählung auswirken. Bei einer zeitlichen Verzögerung von drei Wochen beispielsweise berichteten ehrliche Personen deutlich weniger Details als bei einer Befragung direkt nach dem Ereignis, während die Menge an berichteten Details bei lügenden Personen konstant blieb (Harvey, Vrij, Hope et al., 2017; Harvey, Vrij, Leal et al., 2017). Dies führte dazu, dass lügende und ehrliche Personen bei einer zeitlich verzögerten Befragung ähnlich viele Details berichteten (Harvey, Vrij, Hope et al., 2017; Harvey, Vrij, Leal et al., 2017). Ferner war die Schilderung einer ehrlichen Person bei einer Befragung direkt nach dem Ereignis nur dann detailreicher als jene einer lügenden Person, wenn sich die ehrliche Person der Relevanz des Ereignisses *zum Wahrnehmungszeitpunkt* bewusst war. War dies nicht der Fall, unterschieden sich wahre und gelogene Aussagen nicht in ihrem Detailgehalt (Harvey, Vrij, Leal et al., 2017). Für den Nachprüfbarkeits-Ansatz (Jupe et al., 2020) und die Modellaussage (Harvey et al., 2019) legen erste Studien nahe, dass Effekte nach einer zeitlichen Verzögerung verschwinden oder sich gar umkehren können (siehe 4.3. und 4.4.). Diese Ergebnisse machen deutlich, dass Vernehmer beim Einsatz von detailbezogenen Befragungsstrategien unbedingt berücksichtigen sollten, wie viel Zeit seit dem Ereignis vergangen ist und ob eine unschuldige Person das Ereignis zum Wahrnehmungszeitpunkt als relevant erkannt haben dürfte oder eher nicht.

Von der Gruppe zum Einzelfall und die Idee des „verbal baselinings“

Die Situation in Studien unterscheidet sich aber auch noch in anderer Hinsicht von der Einschätzung eines spezifischen Einzelfalls. So werden in Studien die Ergebnisse über eine große Anzahl von Personen gemittelt, wodurch die Eigenheiten einer einzelnen Person weniger ins Gewicht fallen. Hingegen spielt es beim Einzelfall durchaus eine Rolle, wie sich die Person normalerweise verhält, ob sie also beispielsweise eher eloquent oder aber eher wortkarg ist (siehe auch 2.2. *Vernachlässigung der Baseline*). Zudem werden in Studien jeweils wahre und gelogene Aussagen erhoben, wodurch man einen Vergleichswert erhält. Mit statistischen Modellen kann dann der ideale cut-off Wert ermittelt werden.²⁹ Im Einzelfall stellt sich hingegen die Frage, wann ein Hinweis denn nun stark oder schwach ausgeprägt ist. D.h. beispielsweise, *wie viele* Details muss eine Aussage im Einzelfall konkret enthalten, damit sie als vermutlich wahr eingestuft werden kann?

Einige der neuen Befragungsstrategien erlauben es nun, eine Art Baseline oder Vergleichswert für die Person innerhalb der kritischen Situation zu ermitteln. Konkret geht es darum, zwei Situationen innerhalb der kritischen Befragungsphase zu vergleichen, von

²⁹ Da keine der spezifischen Strategien zu Hinweisen führt, welche eine absolut treffsichere Unterscheidung von Lüge und Wahrheit erlauben würden, geht es beim cut-off Wert um jenen Wert, dessen Anwendung zur höchstmöglichen Trefferquote führt.

denen bekannt ist, dass lügende und ehrliche Personen im Vergleich oftmals unterschiedliche Hinweise zeigen. So kommt es bei lügenden Personen z.B. eher zu Inkonsistenzen zwischen der mündlichen Aussage und der angefertigten Skizze als bei Personen, die die Wahrheit sagen. Werden im Vergleich der beiden Vernehmungsabschnitte primär Hinweise beobachtet, die mit einer Lüge in Verbindung stehen, deutet dies darauf hin, dass der Befragte lügen könnte (und umgekehrt). Vrij (2016) hat für diese Art des Vergleichs den Begriff „*verbal baselining*“ geprägt. Tabelle 2 gibt eine kurze Übersicht über die Vergleichsmöglichkeiten innerhalb einer Aussage für die verschiedenen Befragungsstrategien (für weitere Informationen, siehe Vrij, 2016).

Tabelle 2: Vergleichsmöglichkeiten innerhalb der Person für die spezifischen Befragungsstrategien (nach Vrij, 2016)

| | Was wird verglichen? | Mögliche Hinweise innerhalb der Person |
|--|---|--|
| Unerwartete Fragen | | |
| <i>Rückwärts erzählen</i> | Bericht in chronologischer vs. umgekehrter Reihenfolge | <ul style="list-style-type: none"> • bei wahrer Aussage werden eher zusätzliche (im Sinne von neuen) Informationen erwartet • bei einer Lüge werden eher Inkonsistenzen zwischen den beiden Aussagen erwartet |
| <i>Inhaltsbezogen</i> | Antworten auf erwartete vs. unerwartete Fragen | <ul style="list-style-type: none"> • Antworten auf unerwartete Fragen zu einem ähnlichen Inhalt sind bei Lüge <i>und</i> Wahrheit zumeist weniger detailliert, bei einer Lüge ist der Unterschied jedoch ausgeprägter |
| <i>Modalitätsbezogen</i> | Mündlicher Auskunft vs. Skizze | <ul style="list-style-type: none"> • bei einer Lüge gibt es eher Diskrepanzen zwischen den Informationen in der mündlichen Aussage und der Skizze |
| Anregung zum detaillierteren Erzählen (Modellaussage) | Bericht vor vs. nach dem Vorlegen einer Modellaussage | <ul style="list-style-type: none"> • zusätzliche Details nach der Modellaussage wirken bei einer wahren Aussage plausibler • eine wahre Aussage enthält mehr zusätzliche zentrale und periphere Details, eine Lüge hingegen primär zusätzliche periphere Details • mehr zusätzliche Komplikationen bei einer wahren Aussage |
| Nachprüfbarkeits-Ansatz | Anteil an nachprüfbaren Details am Total (d.h. personenbezogene Unterschiede in der Detailliertheit des Berichts werden berücksichtigt) | <ul style="list-style-type: none"> • der Anteil an nachprüfbaren Details gemessen am Total der berichteten Details (nachprüfbar + nicht nachprüfbar) ist in einer wahren Aussage höher |
| SUE-Technik | Inkonsistenzen innerhalb der Aussage (Anmerkung der Autorin: Verwendung wird wegen aktuell noch unklarer Ergebnislage nicht empfohlen) | <ul style="list-style-type: none"> • mehr Inkonsistenzen innerhalb der Aussage (direkte Widersprüche oder zusätzliche Informationen, um etwas zu erklären) bei schuldigen Personen |

Die Autorin hat in Zusammenhang mit den Vorschlägen zum „verbal baselining“ nach Vrij (2016) zwei Anmerkungen. Erstens sollte die Aussagekraft des Anteils-Maßes aus dem Nachprüfbarkeits-Ansatz nicht überschätzt werden, da das Vorhandensein verifizierbarer Details beispielsweise auch von dem zu berichtenden Ereignis abhängt. Auch weitere Faktoren spielen eine Rolle, wie etwa die Zeit, die zwischen Ereignis und Befragung liegt (die aussagende Person könnte nachprüfbare Details vergessen haben) oder die Motivation der Person, detailliert zu erzählen, um nur ein paar Beispiele zu nennen (siehe auch Nahari, 2019). Möglicherweise wäre hier ein holistischerer Ansatz ähnlich jenem der merkmalsorientierten Inhaltsanalyse (siehe z.B. Volbert, 2008) ein gangbarer Weg. Bei diesem Ansatz werden für die Einschätzung der vorliegenden Hinweise nämlich auch Eigenheiten der Person und des Ereignisses berücksichtigt. Zweitens wurde die Idee, dass es bei schuldigen Personen im Rahmen der SUE-Technik eher zu Inkonsistenzen innerhalb der Aussage kommt, in Zusammenhang mit der sogenannten Evidence-Framing-Matrix³⁰ zwar durch zwei erste Studien unterstützt (Granhag et al., 2013 und 2015), jedoch konnte eine dritte Studie diese Befunde nicht bestätigen (Luke et al., 2013). In dieser dritten Studie wurden Aussagen generell selten geändert und entgegen den theoretischen Annahmen fügten unschuldige Personen ihren Aussagen als Reaktion auf die schrittweisen Enthüllungen sogar mehr zusätzliche Informationen hinzu als schuldige Personen. Aus Sicht der Autorin sind daher weitere Studien erforderlich, bevor eine Empfehlung für die Praxis abgegeben werden kann.

Das „verbal baselining“ löst jedoch nicht alle Probleme. Da der Vergleich innerhalb der Aussage der vernommenen Person stattfindet, ist es einerseits zwar weniger wahrscheinlich, dass z.B. die Veränderung in der Menge an Details – um bei diesem Beispiel zu bleiben – durch personenbezogene Faktoren (z.B. Eloquenz oder Vorbereitung auf die Vernehmung) bedingt ist. Andererseits spielen andere Faktoren aber nach wie vor eine Rolle. Dazu gehören etwa Eigenheiten des Ereignisses selbst (z.B. War es komplex genug, dass es überhaupt viele Details hergibt?) oder die Güte, mit der sich die Person noch an das Ereignis erinnern kann. Deshalb kann die Frage nach dem kritischen cut-off Wert (Ab wie vielen berichteten Details soll eine Aussage als „wahr“ beurteilt werden?), leider auch mit dem „verbal baselining“ nicht beantwortet werden. Dennoch ist das „verbal baselining“ zumindest ein Schritt in die richtige Richtung, da es weniger wahrscheinlich ist, dass die Hinweise auf Eigenheiten der Person zurückzuführen sind und daher eher ein Indikator für den Wahrheitsgehalt der Aussage sein können (Vrij, 2016).

³⁰ Bei dieser Art der Offenlegung der Beweise wird mit der Stärke der Quelle und der Spezifität der Information gearbeitet. Beispielsweise kann mit schwacher Quelle und geringer Spezifität gestartet werden (fand das Delikt in der Innenstadt von Bern statt, könnte eine erste Frage lauten: „Wir wissen, dass Sie in Bern waren. Was sagen Sie dazu?“), bevor dann sukzessive Fragen mit stärkerer Quelle und höherer Spezifität gestellt werden (z.B. „Es gibt Bilder einer Überwachungskamera, die Sie um die Tatzeit herum in unmittelbarer Nähe des Tatorts zeigen. Was sagen Sie dazu?“).

Fazit

Bei der Übertragung der Befunde aus Studien auf den Einzelfall in der Praxis gibt es drei zentrale Herausforderungen: Erstens muss der Einfluss von Unterschieden zwischen der Situation im Labor und der Praxis berücksichtigt werden (z.B. Konsequenzen für vernommene Person, welcher nicht geglaubt wird; Zeitdauer, die zwischen Ereignis und Aussage liegt). Zweitens müssen im Gegensatz zu den Gruppenvergleichen in Studien beim Einzelfall die Eigenheiten der Person bedacht werden, also ob eine Person beispielsweise generell zu detailliertem Erzählen neigt oder nicht. Für einige der spezifischen Befragungsstrategien wird solchen Eigenheiten durch einen Vergleich innerhalb der Person Rechnung getragen („verbal baselining“). Insbesondere für die SUE-Technik gibt es aber aktuell keine befriedigende Möglichkeit eines Vergleichs innerhalb der Person. Das Problem des fehlenden cut-off Werts, d.h. der Information, wann nun ein Hinweis stark genug ausgeprägt ist, damit ein Lüge- oder Wahr-Urteil angezeigt ist, bleibt aber in jedem Fall bestehen. Daher kann der Befrager anhand der Hinweise eben immer nur eine *Einschätzung* vornehmen, welche durchaus

5.3. Welche Strategien sind für einen Einsatz in der Praxis bereit?

Vrij (2018; siehe auch Vrij & Fisher, 2016) hat verschiedene Befragungsstrategien dahingehend analysiert, ob ihr Einsatz in der Praxis empfohlen werden kann. Dazu hat er anhand von 14 Kriterien überprüft, ob die Befunde zu einer Strategie insgesamt hinreichend robust, generalisierbar und wissenschaftlich unumstritten sind. Aufgrund der Fehlerrate von ca. 30% geht es bei diesen Empfehlungen ausschließlich um einen allfälligen Einsatz der Strategien in Vernehmungen und nicht etwa um einen Einsatz als Beweismittel bei Gericht. Vrij (2018) argumentiert, dass eine solche Fehlerrate für Vernehmungen akzeptabel sei, da keine abschließenden Urteile gefällt werden, sondern „nur“ Entscheide zum weiteren Vorgehen im Rahmen der Strafverfolgung. Die Autorin stimmt dieser Einschätzung unter folgendem Vorbehalt zu: Es ist essentiell, dass der Vernehmer sich der Grenzen dessen, was diese Strategien leisten können, bewusst ist und eine allfällige Glaubhaftigkeitseinschätzung nur als *einen Hinweis* betrachtet, für dessen Auftreten es auch alternative Erklärungen zu überprüfen gilt (siehe 5.4.).

Vrij (2018) kam aufgrund seiner Analysen zum Schluss, dass ein Einsatz der Technik des Anregens zum detaillierteren Erzählen (Modellaussage), des Nachprüfbarkeits-Ansatzes sowie der SUE-Technik in der Praxis befürwortet werden kann. Bezüglich des Ansatzes der unerwarteten Fragen hält er fest, dass weitere Studien für eine Empfehlung notwendig seien. Das Rückwärtserzählen schließlich ordnet Vrij den Techniken zu, die die kognitive Belastung *aller* Vernommenen erhöht. Diese Techniken beurteilt er als aktuell nicht für die Praxis einsetzbar, da es Bedenken gibt, dass die durch das Rückwärts erzählen verursachte kognitive Belastung auch ehrliche Personen „überfordern“ könnte, wodurch diese dann ähnliche Hinweise wie lügende Personen zeigen würden. Hier ist also Vorsicht geboten.

Obwohl diverse Strategien die von Vrij und Fisher (2016) und Vrij (2018) aufgestellten Kriterien erfüllen und ihr Einsatz in der Praxis daher gemäß den Autoren empfohlen werden kann, müssen bei ihrem Einsatz einige wichtige Punkte berücksichtigt werden. Diese Punkte

wurden bereits an anderer Stelle ausführlich diskutiert und seien hier nur noch einmal kurz zusammengefasst. Bei der SUE-Technik bedarf es weiterer Untersuchungen zur Frage, ob und wann auch unschuldige Personen lügen und damit den vorhandenen Beweisen widersprechen könnten. Im Zusammenhang mit dem Nachprüfbarkeitsansatz ist nach aktuellem Kenntnisstand eine systematische Analyse der Aussagen für eine Glaubhaftigkeitseinschätzung erforderlich. Und schließlich müssen sowohl beim Nachprüfbarkeits-Ansatz als auch beim Anregen detaillierteren Erzählens berücksichtigt werden, (1) wie lange das Ereignis zurückliegt und (2) inwiefern die Person die Relevanz des Ereignisses zum Zeitpunkt des Geschehens erkannt hat (siehe Abschnitt 5.1). Denn obwohl es im Rahmen einer Vernehmung um eine *Einschätzung* der Glaubhaftigkeit und nicht um ein abschließendes Urteil geht, können Fehleinschätzungen weitreichende Folgen für die weiteren Ermittlungen haben und müssen daher möglichst vermieden werden (siehe 5.4.)

Fazit

Vrij (2018) hat die spezifischen Befragungsstrategien dahingehend analysiert, wie ausgereift und gut erforscht sie sind. Er kam zum Schluss, dass diverse dieser Strategien für einen Einsatz in der Praxis bereit sind, wobei sich diese Empfehlung aber ausschließlich auf Vernehmungen durch Polizei oder Staatsanwaltschaft bezieht. Aber auch bei jenen Strategien, die für einen Einsatz empfohlen werden, gilt es, gewisse Punkte zu beachten, um Fehleinschätzungen möglichst zu vermeiden.

5.4. Berücksichtigen von Alternativerklärungen

Obwohl die Studienbefunde zu den spezifischen Befragungsstrategien teilweise recht vielversprechend sind, liegt die Urteilskorrektheit weit unter 100%. Es ist also damit zu rechnen, dass es auch zu Fehleinschätzungen kommt. Umso wichtiger ist es, dass der Befragte sich der Grenzen dessen, was die Befragungsstrategien leisten können, bewusst ist. Denn sonst läuft er Gefahr, seiner Einschätzung nach dem Einsatz einer solchen Strategie blind zu vertrauen. Obwohl das finale Urteil in einem Fall nicht der Polizei obliegt, sondern dem Gericht, kann eine starke (richtige oder falsche) Überzeugung eines Ermittlers sich auf die weiteren Ermittlungen im Sinne eines Dominoeffekts auswirken. Dabei kann die Einschätzung des Ermittlers dessen Wahrnehmen, Denken und Handeln in Bezug auf andere Tätigkeiten, wie beispielsweise die Beweissicherung, systematisch beeinflussen und verzerren (z.B. Kassin et al., 2013; Oswald & Wyler, 2018).

Um solch systematischen Verzerrungen in den Ermittlungen möglichst entgegenzuwirken, sollten die folgenden beiden Punkte unbedingt beachtet werden. Erstens dürfen eine Befragungsstrategie bzw. die durch sie hervorgebrachten Hinweise nie dazu genutzt werden, eine *abschließende* Einschätzung zum Wahrheitsgehalt einer Aussage oder zur Schuld einer Person zu fällen. Durch den Einsatz der Strategien können lediglich *Hinweise* in die eine oder andere Richtung gewonnen werden. Zweitens müssen insbesondere dann, wenn die Hinweise darauf hindeuten, dass die Aussage nicht wahrheitsgemäß ist, mögliche

Alternativerklärungen in Betracht gezogen und überprüft werden.³¹ Könnte diese Person eine andere Motivation gehabt haben, der Evidenz zu widersprechen, als die, ihre Täterschaft zu verschleiern (z.B. eine unschuldige Person befürchtet, dass ihr nicht geglaubt wird, wenn sie zugibt, am Tatort gewesen zu sein)? Gab es bei dem Ereignis möglicherweise einfach kaum nachprüfbar Details, die die Person hätte berichten können? Liegt das Ereignis weit zurück und/oder war es vielleicht nicht besonders distinktiv oder relevant, sodass die Person eine eher schwache Erinnerung daran hat? Und so weiter. Werden diese beiden Punkte systematisch beherzigt, sollte einem wohl überlegten Einsatz der spezifischen Befragungsstrategien nichts im Wege stehen.

Fazit

Obwohl die spezifischen Befragungsstrategien die korrekte Einschätzung des Wahrheitsgehalts einer Aussage begünstigen können, kann es zu Fehlern kommen. Es ist deshalb von höchster Wichtigkeit, dass der Befrager (1) sich der Grenzen dessen bewusst ist, was mit den neuen Strategien erreicht werden kann, und (2) immer auch mögliche Alternativerklärungen für seine Beobachtungen in Betracht zieht.

5.5. Relevanz von Ton- oder Bild-Ton-Aufzeichnungen von Vernehmungen

Für einen optimalen Einsatz der spezifischen Befragungsstrategien ist eine Ton- oder Bild-Ton-Aufzeichnung der Vernehmung unabdingbar. Bei einem Vernehmungsprotokoll gehen naturgemäß viele Informationen unwiederbringlich verloren. In der Schweiz beispielsweise müssen nur entscheidende Fragen und Antworten wörtlich protokolliert werden (Capus, 2014), und auch in Deutschland besteht in den meisten Fällen³² keine Anforderung, dass es sich um ein *Wortprotokoll* handeln muss (Altenhain, 2015). Die Inhalte einer Aussage werden also oftmals nur selektiv festgehalten, wobei diese Selektion von verschiedenen Faktoren beeinflusst wird, wie etwa von Capus (2014) ausführlich darlegt wird. Ein Schriftprotokoll ist daher keine hinlängliche Grundlage für eine systematische Analyse hinsichtlich des Auftretens inhaltlicher Hinweise. Aber selbst wenn Befragungsstrategien „intuitiv“ genutzt werden sollen, ist eine Aufzeichnung der Vernehmung von ausgesprochen großem Vorteil. Denn es scheint wahrscheinlich, dass der Befrager in der Vernehmungssituation nicht über hinreichende kognitive Ressourcen verfügt, um gleichzeitig die Vernehmung durchzuführen, die Aussagen des Vernommenen bestmöglich

³¹ Dies einerseits deshalb, weil aufgrund einer möglichen Schuldvorannahme (siehe 2.2. *Bestätigungsfehler*) allenfalls nicht alle be- und entlastenden verbalen Hinweise gleichermaßen wahrgenommen und berücksichtigt wurden, und andererseits, weil eine falsch positive Entscheidung (d.h. eine ehrliche Person wird als Lügner fehlklassifiziert) aufgrund eines möglichen Dominoeffekts schwerwiegende Konsequenzen haben kann. Die Überprüfung von Alternativerklärungen für die beobachteten Hinweise ist aber in jedem Fall ratsam.

³² Gemäß der Änderung von §136 Abs. 4 StPO, welche am 01.01.2020 in Kraft getreten ist, gilt, dass Vernehmungen des Beschuldigten audiovisuell aufzuzeichnen sind, wenn „1. dem Verfahren ein vorsätzlich begangenes Tötungsdelikt zugrunde liegt und einer Aufzeichnung weder die äußeren Umstände noch die besondere Dringlichkeit der Vernehmung entgegenstehen oder 2. die schutzwürdigen Interessen von Beschuldigten, die erkennbar unter eingeschränkten geistigen Fähigkeiten oder einer schwerwiegenden seelischen Störung leiden, durch die Aufzeichnung besser gewahrt werden können“. Daher sollte bei vorsätzlichen Tötungsdelikten neu im Normalfall die wörtliche Protokollierung der Aussagen mittels eines Transkripts vorliegen (R. Deckers, persönliche Mitteilung vom 06.06.2020).

zu verarbeiten, die Aussage hinsichtlich der relevanten Hinweise einzuschätzen und sich auch noch mit möglichen Alternativerklärungen für beobachtete Hinweise zu befassen. Eine Aufzeichnung erlaubt es dem Befrager, die Aussage im Anschluss oder zu einem späteren Zeitpunkt nochmals durchzugehen und seine Einschätzung in Ruhe vornehmen oder überprüfen zu können. Ferner erlaubt eine solche Aufzeichnung eine *systematische* Analyse der Aussagen hinsichtlich des Vorhandenseins spezifischer Merkmale, und zwar sowohl durch den Vernehmer im Anschluss an die Vernehmung als auch durch eine spätere Instanz. Schließlich würde eine Aufzeichnung den Befrager während der Vernehmung zusätzlich kognitiv entlasten, wenn dadurch das gleichzeitige Protokollieren entfallen würde. Die freigewordenen Ressourcen stünden dann beispielsweise für die Befragungsgestaltung und den Beziehungsaufbau zur Verfügung. Die gesetzlichen Grundlagen für eine Aufzeichnung der Vernehmung sind sowohl in Deutschland als auch in der Schweiz gegeben.³³ Es bleibt nur zu hoffen, dass sich eine flächendeckende Dokumentation von Vernehmungen im Sinne von Ton- oder Bild-Ton-Aufzeichnungen möglichst bald durchsetzen wird. Ein erster Schritt in diese Richtung ist die in Deutschland seit dem 01.01.2020 in Kraft getretene geänderte Vorschrift des § 136 Abs. 4 StPO (vgl. Fußnote 32). Dieser gesetzgeberische Fortschritt legt durchaus nahe, die ansonsten hinsichtlich der Bild-Ton-Aufzeichnung von Beschuldigtenvernehmungen geltende „Kannvorschrift“ bei schwerwiegenden Vorwürfen jenseits der Tötungsdelikte als „Sollvorschrift“ zu interpretieren (R. Deckers, persönliche Mitteilung vom 06.06.2020).

Fazit

Der Einsatz spezifischer Befragungsstrategien erhöht die kognitive Belastung des Vernehmers während der Vernehmung zusätzlich. Das Durchführen einer Vernehmung, bei der gleichzeitig die Prinzipien des informationssammelnden Befragungsstils befolgt, spezifische Befragungsstrategien umgesetzt und auf mögliche Hinweise geachtet werden soll, scheint nicht möglich, ohne dass Qualitätsabstiche bei der einen oder anderen Aufgabe gemacht werden müssen. Eine Ton- oder Bild-Ton-Aufzeichnung der Vernehmung ist daher zu empfehlen.

6. Abschließendes Fazit

Die psychologische Forschung hat diverse vielversprechende Befragungsstrategien hervorgebracht, von denen einige durchaus zum jetzigen Zeitpunkt für einen Einsatz in der Praxis in Betracht gezogen werden können. Diese Strategien können im Rahmen des informationssammelnden Befragungsstil eingesetzt werden, welcher bereits für sich genommen positive Auswirkungen auf den Erkenntnisgewinn in Vernehmungen hat. So erhöht er oftmals die Kooperation bei Aussagenden, schafft gute Bedingungen für einen Erinnerungsabruf bei aussagewilligen Personen und verringert das Risiko von falschen Geständnissen. Ferner kann ein konsequentes Verfolgen einer möglichst offenen und unvoreingenommenen Herangehensweise an die Befragung den Befrager selbst in der wichtigen Bemühung unterstützen, stetig auch alternative Erklärungsmöglichkeiten für Informationen und Hinweise zu überprüfen und zu berücksichtigen.

³³ Deutschland: § 136 Abs. 4 Satz 1 StPO; Schweiz: Art. 76 Abs. 4 StPO.

Die durch eine Befragungsstrategie hervorgebrachten Hinweise können oftmals „intuitiv“ vom Befrager genutzt werden, obwohl bei einer systematischeren Herangehensweise vermutlich noch etwas zuverlässigere Einschätzungen möglich wären. Befrager müssen sich aber auch im Klaren darüber sein, dass alle spezifischen Befragungsstrategien mit einer Fehlerrate um die 30% behaftet sind. Einige Forscher argumentieren, dass solche Fehlerraten in Vernehmungen, bei denen es nicht um ein abschließendes Urteil, sondern um das weitere Vorgehen bei den Ermittlungen geht, akzeptabel ist. Grundsätzlich befürwortet auch die Autorin die Nutzung von spezifischen Befragungsansätzen in einem solchen Kontext, jedoch nur unter bestimmten Bedingungen. Insbesondere darf eine Einschätzung des Wahrheitsgehalts einer Aussage, welche auf Basis der Nutzung einer spezifischen Befragungsstrategie erfolgte, nur als das gesehen werden, was sie ist: Ein Hinweis in eine bestimmte Richtung. Dieser Hinweis muss zwingend im Kontext anderer vorliegender Informationen interpretiert werden und es gilt, alternative Erklärungen für das Auftreten von bestimmten Hinweisen kritisch zu prüfen. Denn eine Fehleinschätzung kann sich nicht nur auf den weiteren Verlauf der Befragung auswirken, sondern auch die weiteren Ermittlungen verzerren, da diese Einschätzung das weitere Denken und Handeln des Ermittlers (z.B. soll eine bestimmte Spur weiterverfolgt werden?) maßgeblich beeinflussen kann. Um das Risiko von Fehleinschätzungen zu reduzieren, kann sich der Vernehmer jeweils folgende Fragen stellen:

1. Wie gut könnte sich eine die Wahrheit sagende Person vermutlich an das Ereignis erinnern?
 - i. Wie viel Zeit ist seit dem Ereignis vergangen?
 - ii. Wäre sich eine ehrliche/unschuldige Person der Relevanz des Ereignisses bereits zum *Wahrnehmungszeitpunkt* bewusst gewesen?
2. Scheint sich die Person in der Vernehmung einigermaßen wohl zu fühlen, d.h., scheinen gute Bedingungen für eine Person vorzuliegen, die sich bestmöglich erinnern will?
3. Gibt es andere Gründe, weshalb eine die Wahrheit sagende Person mit dem Erinnern Schwierigkeiten haben könnte (z.B. eingeschränktes kognitives Funktionsniveau)?
4. Welche alternativen Erklärungsmöglichkeiten gibt es für vorliegende Täuschungshinweise (z.B. Eigenheiten des Ereignisses oder Motivation der Person) und könnte eine davon auf den aktuellen Fall zutreffen?

Die Fragen 1-3 sind wichtig, da bei vielen der spezifischen Befragungsstrategien der Detailgehalt einer Aussage in der einen oder anderen Form von Bedeutung ist. Liegen nun Umstände vor, die die Erinnerung an das Ereignis oder aber die Fähigkeit zum Erinnerungsabruf in der Vernehmungssituation beeinträchtigen, kann es auch bei ehrlichen Personen zu einem eher niedrigen Detaillierungsgrad kommen. Frage 4 ist u.a. in Bezug auf die SUE-Technik wichtig, da nicht ausgeschlossen werden kann, dass auch eine unschuldige Person den Beweisen unter bestimmten Umständen widerspricht.

In Vernehmungen wird Befragern kognitiv sehr viel abverlangt. Dies gilt umso mehr, wenn zusätzlich spezifische Befragungsstrategien eingesetzt und allfällig auftretende Hinweise analysiert werden sollen. Daher scheint für einen sinnvollen Einsatz solcher Befragungsstrategien eine Ton- oder Bild-Ton-Aufzeichnung der Vernehmung fast

unabdingbar. Schließlich sollte man Vorsicht walten lassen was den simultanen Einsatz mehrerer spezifischer Befragungsstrategien anbelangt. Der Einsatz zu vieler oder aber ungeeigneter Strategien in einer Vernehmung könnte auch ehrliche Personen überfordern oder aber ihre Motivation zur Mitarbeit reduzieren, wenn sie sich mit allzu vielen Forderungen konfrontiert sehen. Obwohl es grundsätzlich möglich ist, verschiedene Befragungsstrategien untereinander zu kombinieren, sollte eine solche Kombination also immer sorgfältig überlegt sein.

Abschließend sei darauf hingewiesen, dass einige der Befragungsstrategien auch einen allfälligen Erkenntnisgewinn begünstigen, der über die Möglichkeit einer verbesserten Glaubhaftigkeitseinschätzung in der Vernehmung hinausgeht. So können zusätzliche nachprüfbare Details, aber auch anderweitige, zusätzlich berichtete Informationen nicht nur zu neuen Spuren in den Ermittlungen führen, sondern auch für die Beurteilung des Falls durch das Gericht von Bedeutung sein.

Literaturverzeichnis

- Akehurst, L., Köhnken, G., Vrij, A. & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology*, 10(6), 461–471.
- Altenhain, K. (2015). Dokumentationspflicht im Ermittlungsverfahren. Warum eigentlich nicht? *Zeitschrift für Internationale Strafrechtsdogmatik*, 5, 269–282.
- Bensi, L., Nori, R., Gambetti, E. & Giusberti, F. (2011). The Enhanced Cognitive Interview: A study on the efficacy of shortened variants and single techniques. *Journal of Cognitive Psychology*, 23(3), 311–321.
- Blair, J. P. & Kooi, B. (2004). The gap between training and research in detection of deception. *International Journal of Police Science & Management*, 6(2), 77–83.
- Bogaard, G., Meijer, E. H. & Vrij, A. (2014). Using an example statement increases information but does not increase accuracy of CBCA, RM, and SCAN. *Journal of Investigative Psychology and Offender Profiling*, 11(2), 151–163.
- Bogaard, G., Meijer, E. H. & Van der Plas, I. (2020). A model statement does not enhance the verifiability approach. *Applied Cognitive Psychology*, 34(1), 96–105.
- Böhme, T. (2018). *Das strafgerichtliche Fehlurteil – Systemimmanenz oder vermeidbares Unrecht? Eine Untersuchung zu den Ursachen von Fehlurteilen im Strafprozess und den Möglichkeiten ihrer Vermeidung. Schriftenreihe Deutsche Strafverteidiger e.V. (Band 46). Nomos.*
- Bond, C. F. & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214–234.
- Brockmann, C. (2008). Beschuldigtenvernehmung. In R. Volbert & M. Steller (Hrsg.), *Handbuch der Rechtspsychologie* (S. 244-252). Hogrefe.
- Bull, R. & Milne, B. (2004). Attempts to improve the police interviewing of suspects. In D. Lassiter (Hrsg.), *Interrogations, confessions, and entrapment* (S. 181-196). Kluwer Academic/Plenum Publishers.
- Bull, R. & Soukara, S. (2010). Four studies of what really happens in police interviews. In G. D. Lassiter & C. A. Meissner (Hrsg.), *Interrogations and false confessions: Current research, practice, and policy recommendations* (S. 81-95). American Psychological Association.
- Capus, N. (2014). «Ich ermördere Dich. Es gibt keine Gesetze für Vito. Vitogesetze.» Theorie und Empirie zur Herstellung von Schriftprotokollen in Strafverfahren. *Justice – Justiz – Giustizia*, 3, 1–14.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Academic Press.

- Crozier, W. E., Strange, D. & Loftus, E. F. (2017). Memory errors in alibi generation: How an alibi can turn against us. *Behavioral Sciences and the Law*, 35(1), 6–17.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K. & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74–118.
- Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. Duncker & Humblot.
- Evans, J. R., Meissner, C. A., Ross, A. B., Houston, K. A., Russano, M. B. & Horgan, A. J. (2013). Obtaining guilty knowledge in human intelligence interrogations: Comparing accusatorial and information-gathering approaches with a novel experimental paradigm. *Journal of Applied Research in Memory and Cognition*, 2(2), 83–88.
- Fincham, F. & Hewstone, M. (2003). Attributionstheorie und -forschung – Von den Grundlagen zur Anwendung. In W. Stroebe, K. Jonas & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (4. Aufl., S. 215-264). Springer.
- Fisher, R. P. & Geiselman, R. E. (1992). *Memory-enhancing techniques in investigative interviewing: The cognitive interview*. C. C. Thomas.
- Geiselman, R. E. & Callot, R. (1990). Reverse versus forward recall of script-based texts. *Applied Cognitive Psychology*, 4(2), 141–144.
- Gilbert, J. A. E. & Fisher, R. P. (2006). The effects of varied retrieval cues on reminiscence in eyewitness memory. *Applied Cognitive Psychology*, 20(6), 723–739.
- Gilovich, T., Savitsky, K. & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, 75(2), 32–346.
- Global Deception Research Team (2006). A world of lies. *Journal of Cross-Cultural Psychology*, 37(1), 60–74.
- Granhag, P. A. & Hartwig, M. (2015). The Strategic Use of Evidence Technique: A conceptual overview. In P. A. Granhag, A. Vrij & B. Verschuere (Hrsg.), *Detecting deception: Current challenges and cognitive approaches* (S. 232-251). Wiley Blackwell.
- Granhag, P. A., Rangmar, J. & Strömwall, L. A. (2015). Small cells of suspects: Eliciting cues to deception by strategic interviewing. *Journal of Investigative Psychology and Offender Profiling*, 12(2), 127–141.
- Granhag, P. A., Strömwall, L. A., Willén, R. M. & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal and Criminological Psychology*, 18(2), 341–355.
- Grice, P. H. (1993). Logik und Konversation. In G. Meggle (Hrsg.), *Handlung, Kommunikation, Bedeutung* (S. 243-265). Suhrkamp.

- Haas, H. & Ill, C. (2013). Gesprächsführungstechniken in der Einvernahme. *Forum poenale*, 0, 2–27.
- Habschick, K. (2012). *Erfolgreich Vernehmen. Kompetenz in der Kommunikations-, Gesprächs- und Vernehmungspraxis*. Kriminalistik Verlag.
- Hartwig, M. (2011). Methods in deception detection research. In B. Rosenfeld & S. D. Penrod (Hrsg.), *Research methods in forensic psychology* (S. 136-155). Wiley.
- Hartwig, M. & Bond, C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgements. *Psychological Bulletin*, 137(4), 643–659.
- Hartwig, M. & Bond, C. F. (2014). Lie detection from multiple cues: A meta-analysis. *Applied Cognitive Psychology*, 28(5), 661–676.
- Hartwig, M., Granhag, P. A. & Luke, T. (2014). Strategic use of evidence during investigative interviews. The state of the science. In D. C. Raskin, C. R. Honts & J. C. Kircher (Hrsg.), *Credibility assessment: Scientific research and applications* (S. 1-36). Academic Press.
- Hartwig, M., Granhag, P. A. & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, 13(2), 213–227.
- Hartwig, M., Granhag, P. A., Strömwall, L. A. & Kronkvist, O. (2006). Strategic Use of Evidence during police interviews: When training to detect deception works. *Law and Human Behavior*, 30, 603–619.
- Hartwig, M., Granhag, P. A., Strömwall, L. A. & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law and Human Behavior*, 29, 469–484.
- Hartwig, M., Granhag, P. A., Strömwall, L. A., Wolf, A. G., Vrij, A. & Roos af Hjelmsäter, E. (2011). Detecting deception in suspects: Verbal cues as a function of interview strategy. *Psychology, Crime & Law*, 17(7), 643–656.
- Harvey, A. C., Vrij, A., Hope, L., Leal, S. & Mann, S. (2017). A stability bias effect among deceivers. *Law and Human Behavior*, 41, 519–529.
- Harvey, A. C., Vrij, A., Leal, S., Hope, L. & Mann, S. (2017). Deception and decay: Verbal lie detection as a function of delay and encoding quality. *Journal of Applied Research in Memory and Cognition*, 6(3), 306–318.
- Harvey, A. C., Vrij, A., Hope, L., Leal, S. & Mann, S. (2019). Amplifying deceivers' flawed metacognition: Encouraging disclosures after delays with a model statement. *Acta Psychologica*, 200, Artikel 102935.
- Hermanutz, M. & Litzcke, S. M. (2009). Polizeiliche Vernehmungen – Bestandsaufnahme. In M. Hermanutz & S. M. Litzcke (Hrsg.), *Vernehmung in Theorie und Praxis: Wahrheit – Irrtum – Lüge* (S. 69-85). Boorberg.

- Hill, C., Memon, A., & McGeorge, P. (2008). The role of confirmation bias in suspect interviews: A systematic evaluation. *Legal and Criminological Psychology, 13*(2), 357–371.
- Horgan, A. J., Russano, M. B., Meissner, C. A. & Evans, J. R. (2012). Minimization and maximization techniques: Assessing the perceived consequences of confessing and confession diagnosticity. *Psychology, Crime & Law, 18*(1), 65–78.
- Inbau, F. E., Reid, J. E., Buckley, J. P. & Jayne, B. C. (2013). *Criminal interrogation and confessions* (5. Ausg.). Jones & Bartlett Learning.
- Johnson, M. K. (2006). Memory and reality. *American Psychologist, 61*(8), 760–771.
- Jordan, S. & Hartwig, M. (2013). On the phenomenology of innocence: The role of Belief in a Just World. *Psychiatry, Psychology and Law, 20*(5), 749–760.
- Jordan, S., Hartwig, M., Wallace, B., Dawson, E. & Xihani, A. (2012). Early versus late disclosure of evidence: Effects on verbal cues to deception, confessions, and lie catchers' accuracy. *Journal of Investigative Psychology and Offender Profiling, 9*(1), 1–12.
- Jupe, L. M., Vrij, A., Leal, S. & Nahari, G. (2020). Fading lies: Applying the verifiability approach after a period of delay. *Psychology, Crime & Law, 26*(4), 327–342.
- Kahana, M. J. (1996). Associate retrieval processes in free recall. *Memory & Cognition, 24*, 103–109.
- Kassin, S. M. (2005). On the psychology of confessions. Does innocent put innocents at risk? *American Psychologist, 60*(3), 215–228.
- Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A. & Redlich, A. D. (2010). Police-induced confessions: Risk factors and recommendations. *Law and Human Behavior, 34*, 3–38.
- Kassin, S. M., Dror, I. E. & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition, 2*(1), 42–52.
- Kassin, S. M. & Fong, C. T. (1999). 'I'm innocent!' Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior, 23*, 499–516.
- Kassin, S. M., Goldstein, C. C. & Savitsky, K. (2003). Behavioral confirmation in the interrogation room: On the dangers of presuming guilt. *Law and Human Behavior, 27*, 187–203.
- Kassin, S. M. & Gudjonsson, G. H. (2004). The psychology of confessions A review of the literature and issues. *Psychological Science in the Public Interest, 5*(2), 33–67.

- Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A.-M. & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior, 31*, 381–400.
- Lancaster, G. L. J., Vrij, A., Hope, L. & Waller, B. (2013). Sorting the liars from the truth tellers: The benefits of asking unanticipated questions on lie detection. *Applied Cognitive Psychology, 27*(1), 107–114.
- Leal, S., Vrij, A., Deeb, H. & Jupe, L. (2018). Using the model statement to elicit verbal differences between truth tellers and liars: The benefit of examining core and peripheral details. *Journal of Applied Research in Memory and Cognition, 7*(4), 610–617.
- Leal, S., Vrij, A., Warmelink, L., Vernham, Z. & Fisher, R. P. (2015). You cannot hide your telephone lies: Providing a model statement as an aid to detect deception in insurance telephone calls. *Legal and Criminological Psychology, 20*(1), 129–146.
- Leins, D. A., Fisher, R. P. & Vrij, A. (2012). Drawing on liars' lack of cognitive flexibility: Detecting deception through varying report modes. *Applied Cognitive Psychology, 26*(4), 601–607.
- Leins, D. A., Fisher, R. P., Vrij, A., Leal, S. & Mann, S. (2011). Using sketch-drawing to induce inconsistency in liars. *Legal and Criminological Psychology, 16*(2), 253–265.
- Lerner, M. J. (1980). *The belief in a just world*. Plenum.
- Litzcke, S. M. & Klossek, A. (2009). Glaubhaftigkeitsattribution – Wahrheits- und Lügenstereotype. In M. Hermanutz & S. M. Litzcke (Hrsg.), *Vernehmung in Theorie und Praxis: Wahrheit – Irrtum – Lüge* (S. 199–212). Bloorberg.
- Luke, T. J., Hartwig, M., Brimbal, L., Chan, G., Jordan, S., Joseph, E., Osborne, J. & Granhag, P. A. (2013). Interviewing to elicit cues to deception: Improving Strategic Use of Evidence with General-To-Specific Framing of Evidence. *Journal of Police and Criminal Psychology, 28*, 54–62.
- Luke, T. J., Hartwig, M., Joseph, E., Brimbal, L., Chan, G., Dawson, E., Jordan, S., Donovan, P. & Granhag, P. A. (2016). Training in the Strategic Use of Evidence technique: Improving deception detection accuracy of American law enforcement officers. *Journal of Police and Criminal Psychology, 31*, 270–278.
- Luke, T. J., Hartwig, M., Shamash, B. & Granhag, P. A. (2016). Countermeasures against the Strategic Use of Evidence Technique: Effects on suspects' strategies. *Journal of Investigative Psychology and Offender Profiling, 13*(2), 131–47.
- Meissner, C. A., Redlich, A. D., Michael, S. W., Evans, J. R., Camilletti, C. R., Bhatt, S. & Brandon, S. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *Journal of Experimental Criminology, 10*, 459–486.

- Miller, J. C., Redlich, A. D. & Kelly, C. H. (2018). Accusatorial and information-gathering interview and interrogation methods: A multi-country comparison. *Psychology, Crime & Law*, 24, 935–956.
- Milne, B., Shaw, G. & Bull, R. (2007). Investigative interviewing: The role of research. In D. Carson, R. Milne, F. Pakes, K. Shalev & A. Shawyer (Hrsg.), *Applying psychology to criminal justice* (S. 65-80). Wiley.
- Moston, S. & Engelberg, T. (1993). Police questioning techniques in tape recorded interviews with criminal suspects. *Policing and Society*, 3(3), 223–237.
- Nahari, G. (2019). Verifiability approach. Applications in different judgmental settings. In T. Doncan-Morgan (Hrsg.), *The Palgrave Handbook of Deceptive Communication* (S. 213–226). Palgrave Macmillan.
- Nahari, G., Vrij, A. & Fisher, R. P. (2012). Does the truth come out in the writing? SCAN as a lie detection tool. *Law and Human Behavior*, 36, 68–76.
- Nahari, G., Vrij, A. & Fisher, R. P. (2014a). Exploiting liars' verbal strategies by examining unverifiable details. *Legal and Criminological Psychology*, 19(2), 227–239.
- Nahari, G., Vrij, A. & Fisher, R. P. (2014b). The Verifiability Approach: Countermeasures facilitate its ability to discriminate between truths and lies. *Applied Cognitive Psychology*, 28(1), 122–128.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231–259.
- Oswald, M. E. & Grosjean, S. (2004). Confirmation bias. In R. F. Pohl (Hrsg.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory*. Psychology Press.
- Oswald, M. E. & Wyler, H. (2018). Fallstricke auf dem Weg zur «richtigen» Entscheidung im Strafrecht: Eine Analyse aus psychologischer Sicht. In S. Barton, M. Dubelaar, R. Köbel & M. Lindemann (Hrsg.), *«Vom hochgemuten, voreiligen Griff nach der Wahrheit»: Fehlurteile im Strafprozess* (S. 103-132). Nomos.
- Palena, N., Caso, L., Vrij, A. & Orthey, R. (2018). Detecting deception through small talk and comparable truth baselines. *Journal of Investigative Psychology and Offender Profiling*, 15(2), 124–132.
- Porter, C. N. & Salvaneli, G. (2020). Eliciting information and cues to deception using a model statement: Examining the effect of presentation modality. *Journal of Investigative Psychology and Offender Profiling*, 1–17. Advance online publication. <https://doi.org/10.1002/jip.1541>

- Porter, S., England, L., Juodis, M., ten Brinke, L. & Wilson, K. (2008). Is the face a window to the soul? Investigation of the accuracy of intuitive judgments of the trustworthiness of human faces. *Canadian Journal of Behavioural Science*, 40(3), 171–177.
- Porter, S., ten Brinke, L. & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, 16(6), 477–491.
- Rassin, E. (2010). Blindness to alternative scenarios in evidence evaluation. *Journal of Investigative Psychology and Offender Profiling*, 7(2), 153–163.
- Shaw, D. J., Vrij, A., Leal, A., Mann, S., Hillman, J., Granhag, P. A. & Fisher, R. P. (2013). Expect the unexpected? Variations in question type elicit cues to deception in joint interviewer contexts. *Applied Cognitive Psychology*, 27(3), 336–343.
- Shields, G. S., Sazma, M. A., McCullough, A. M. & Yonelinas, A. P. (2017). The effects of acute stress on episodic memory: A meta-analysis and integrative review. *Psychological Bulletin*, 143(6), 636–675.
- Steller, M. & Köhnken, G. (1989). Criteria-based statement analysis. In D. C. Raskin (Hrsg.), *Psychological methods in criminal investigation and evidence* (S. 217–245). Springer.
- Strömwall, L. A. & Granhag, P. A. (2003). How to detect deception? Arresting the beliefs of police officers, prosecutors and judges. *Psychology, Crime & Law*, 9(1), 19–36.
- Unsworth, N. & Spillers, G. J. (2010). Variation in working memory capacity and episodic recall: The contributions of strategic encoding and contextual retrieval. *Psychonomic Bulletin & Review*, 17, 200–205.
- Vernham, Z., Vrij, A., Nahari, G., Leal, S., Mann, S., Satchell, L. & Robin, O. (2020). Applying the verifiability approach to deception detection in alibi witness situations. *Acta Psychologica*, 204, Artikel 103020.
- Volbert, R. (2008). Glaubhaftigkeitsbegutachtung – mehr als Merkmalsorientierte Inhaltsanalyse. *Forensische Psychiatrie, Psychologie, Kriminologie*, 2, 12–19.
- Volbert, R. (2012). Falsche Geständnisse – Häufigkeit, Entstehungsbedingungen, Beurteilung. In R. Egg (Hrsg.), *Psychologisch-psychiatrische Begutachtung in der Strafjustiz* (S. 69-90). Eigenverlag Kriminologische Zentralstelle e.V.
- Volbert, R. (2013). Falsche Geständnisse. Über die möglichen Auswirkungen von Voreinstellungen, Vernehmung und Verständigung. *Forensisch Psychiatrie, Psychologie, Kriminologie*, 7, 230–239.
- Volbert, R. & May, L. (2016). Falsche Geständnisse in polizeilichen Vernehmungen: Vernehmungsfehler oder immanente Gefahr? *Recht und Psychiatrie*, 34, 4–10.
- Vrij, A. (2008a). *Detecting lies and deceit. Pitfalls and opportunities* (2. Aufl.). Wiley.

- Vrij, A. (2008b). Nonverbal dominance versus verbal accuracy in lie detection. A plea to change police practice. *Criminal Justice and Behavior*, 35, 1323–1336.
- Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Hrsg.), *Detecting deception: Current challenges and cognitive approaches* (S. 205-229). Wiley-Blackwell.
- Vrij, A. (2016). Baseline as a lie detection method. *Applied Cognitive Psychology*, 30(6), 1112–1119.
- Vrij, A. (2018). Verbal lie detection tools from an applied perspective. In J. P. Rosenfeld (Hrsg.), *Detecting concealed information and deception: Recent developments* (297-327). Academic Press.
- Vrij, A. & Fisher, R. P. (2016). Which lie detection tools are ready for use in the Criminal Justice System? *Journal of Applied Research in Memory and Cognition*, 5(3), 302–307.
- Vrij, A., Fisher, R. P. & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, 22(1), 1–21.
- Vrij, A., Hope, L. & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 129–136.
- Vrij, A., Leal, S. & Fisher, R. P. (2018). Verbal deception and the model statement as a lie detection tool. *Frontiers in Psychiatry*, 9, Artikel 492.
- Vrij, A., Leal, S., Granhag, P. A., Mann, S., Fisher, R. P., Hillman, J. & Sperry, K. (2009). Outsmarting the liars: The benefit of asking unanticipated questions. *Law and Human Behavior*, 33, 159–166.
- Vrij, A., Leal, S., Jupe, L. & Harvey, A. (2018). Within-subjects verbal lie detection measures: A comparison between total detail and proportion of complications. *Legal and Criminological Psychology*, 23(2), 265–279.
- Vrij, A., Leal, S., Mann, S., Dalton, G., Jo, E., Shaboltas, A., Khaleeva, M., Granskaya, J. & Houston, K. (2017). Using the model statement to elicit information and cues to deceit in interpreter-based interviews. *Acta Psychologica*, 177, 44–53.
- Vrij, A., Leal, S., Mann, S. & Fisher, R. (2012). Imposing cognitive load to elicit cues to deceit: Inducing the reverse order technique naturally. *Psychology, Crime & Law*, 18(6), 579–594.
- Vrij, A., Mann, S. & Fisher, R. P. (2006). An empirical test of the Behaviour Analysis Interview. *Law and Human Behavior*, 30, 329–345.

- Vrij, A., Mann, S. A., Fisher, R. P., Leal, S., Milne, R. & Bull, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, 253–265.
- Vrij, A., Mann, S., Kirsten, S. & Fisher, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior*, 31, 499–518.
- Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan C. A. III & Kleinman, S. M. (2017). Psychological perspectives on interrogation. *Perspectives on Psychological Science*, 12(6), 927–955.
- Vrij, A. & Nahari, G. (2018). The verifiability approach. In J. Dickinson, N. Schreiber Compo, R. Carol & M. McCauley (Hrsg.), *Evidence-based investigative interviewing* (116-133). Routledge.