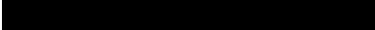


Bachelorarbeit

Die Adaption von Methoden der Human Reliability Analyse in der Energiewirtschaft

Eine Übersichtsarbeit für die DNV Energy
Systems Germany GmbH

Vorgelegt am 30.08.2024 von
Markus Gutermuth


Studiengang Rettungsingenieurwesen

1. Gutachter: Prof. Dr. Marc Schütte
2. Gutachter: Marcus Tatulinski, M. Sc.

**Hochschule für Angewandte
Wissenschaften Hamburg**
Fakultät Life Sciences
Ulmenliet 20

In Zusammenarbeit mit



Abstract

Im Gegensatz zur Nuklearindustrie ist die Human Reliability Analyse (HRA) in der Öl- und Gasindustrie noch kein etabliertes Werkzeug im Risikomanagement. Es fehlt sowohl an regulatorischen Vorgaben als auch an praktischer Erfahrung in der Anwendung. Allerdings wurden erste Methoden entwickelt, die speziell auf die besonderen Anforderungen dieser Branche, wie z. B. die größere Vielfalt an Standortbedingungen, Anlagendesigns und Prozessen, zugeschnitten sind. Basierend auf einer strukturierten Literaturrecherche wird untersucht, wie bestehende HRA-Methoden angepasst werden, um diese industriespezifischen Faktoren zu berücksichtigen. Diese Anpassungen betreffen insbesondere die Modifikation der Performance Shaping Factors. Die Analyse zeigt jedoch, dass viele der vorgenommenen Änderungen vor allem mit der Verbesserung der Anwenderfreundlichkeit begründet wurden und weniger mit spezifischen Anforderungen der Industrie. Darüber hinaus werden Ansätze vorgestellt, die darauf abzielen, Schwächen wie die starke Abhängigkeit von der Expertise der Analysten zu kompensieren, insbesondere durch die Integration empirischer Daten und den Einsatz von Bayes'schen Netzen.

Inhaltsverzeichnis

| | | |
|------|--|----|
| I. | Abkürzungsverzeichnis | iv |
| II. | Abbildungsverzeichnis | v |
| III. | Tabellenverzeichnis..... | vi |
| 1. | Einleitung | 1 |
| 2. | Theoretischer Hintergrund | 3 |
| 2.1 | Sicherheit und der Faktor Mensch..... | 3 |
| 2.2 | Abgrenzung HRA und Human Factor | 6 |
| 2.3 | Allgemeiner Prozess der HRA..... | 7 |
| 2.4 | Historische Entwicklung und Klassifizierung | 9 |
| 2.5 | Trends und Herausforderungen | 11 |
| 2.6 | Regulatorische Rahmenbedingungen und Standards | 13 |
| 3. | Methode | 17 |
| 4. | Ergebnisse | 19 |
| 4.1 | Deskriptive Darstellung | 19 |
| 4.2 | Inhaltliche Darstellung | 19 |
| 5. | Diskussion | 27 |
| 5.1 | Diskussion der Methode | 27 |
| 5.2 | Diskussion der Ergebnisse | 27 |
| | Literaturverzeichnis | 30 |

I. Abkürzungsverzeichnis

| | |
|-------------------|---|
| APOA | Assessed Proportion of Affect |
| ATHEANA | A Technique for Human Error Analysis |
| BBergG | Bundesberggesetz |
| BImSchG | Bundes-Immissionsschutzgesetz |
| BN | Bayes'sches Netz |
| CFM | Crew Failure Modes |
| CREAM | Cognitive Reliability Error Analysis Method |
| EOC | Error of Commission |
| EOO | Error of Omission |
| EPC | Error Producing Condition |
| EPRI | Electric Power Research Institute |
| GT _{HEP} | Generic Task Human Error Probability |
| GTT | Generic Task Type |
| HAZID | Hazard Identification |
| HAZOP | Hazard and Operability |
| HEART | Human Error Assessment and Reduction Technique |
| HEI | Human Error Identification |
| HEP | Human Error Probability |
| HF | Human Factor |
| HFE | Human Failure Events |
| HMI | Human Machine Interface |
| HRA | Human Reliability Analysis |
| HTA | Hierarchical Task Analysis |
| IAEA | International Atomic Energy Agency |
| ISO | International Organization for Standardization |
| KAS | Kommission für Anlagensicherheit |
| LOPA | Layer of Protection Analysis |
| MERMOS | Méthode d'Évaluation de la Réalisation des Missions Opérationnelles de Sûreté |
| NHEP | Nominal Human Error Probability |
| OffshoreBergV | Bergverordnung für das Gebiet der Küstengewässer und des Festlandssockels |
| PHEA | Predictive Human Error Analysis |
| PIF | Performance Influencing Factors |
| PSA | Probabilistic Safety Assessment |
| PSF | Performance Shaping Factors |
| QRA | Quantitative Risk Assessment |
| SHERPA | Systematic Human Error Reduction and Prediction Approach |
| SIL | Safety Integrity Level |
| SPAR-H | Standardized Plant Analysis Risk-Human Reliability Analysis |
| THERP | Technique for Human Error Rate Prediction |
| U.S. NRC | U.S. Nuclear Regulatory Commission |
| 12. BImSchV | Zwölfte Verordnung zur Durchführung des Bundes-Immissionsschutzgesetzes (Störfall-Verordnung) |

II. Abbildungsverzeichnis

| | |
|--|----|
| Abbildung 1: Fehlerklassifikation nach Reason | 4 |
| Abbildung 2: Schematische Darstellung des HRA-Prozesses | 7 |
| Abbildung 3: HRA-Methoden nach Finanzierungsquellen | 11 |
| Abbildung 4: Prozess der strukturierten Literaturrecherche | 18 |
| Abbildung 5: Bayes'sches Netzes..... | 21 |
| Abbildung 6: Exemplarischer Fehlerbaum einer Gasleckage | 24 |
| Abbildung 7: Ansätze für die Definition von HFE | 26 |

III. Tabellenverzeichnis

| | |
|--|----|
| Tabelle 1: Auszug aus der PHEA Error Taxonomy | 8 |
| Tabelle 2: Generic Task Types in HEART | 11 |
| Tabelle 3: Finale Auswahl der Literatur | 19 |
| Tabelle 4: Gegenüberstellung des PSF HMI aus SPAR-H und Petro-HRA..... | 23 |
| Tabelle 5: Phoenix-PRO Crew Failure Modes | 25 |
| Tabelle 6: Phoenix-PRO Performance Influencing Factors..... | 25 |

1. Einleitung

Die Energiewirtschaft ist ein elementarer Wirtschaftszweig unserer modernen Gesellschaft und gilt in Deutschland aufgrund unserer hohen Energieabhängigkeit auch als kritische Infrastruktur. Die Energiewirtschaft zeichnet sich durch eine Vielzahl von Energieträgern und Technologien aus. Auf viele dieser Technologien treffen die von Charles Perrow im Rahmen der *Normal Accident Theory* definierten Eigenschaften eines für Systemunfälle anfälligen soziotechnischen Systems zu. Das heißt, es handelt sich um komplex und eng gekoppelte Systeme, von denen im Falle einer Fehlfunktion ein hohes Gefährdungs- und Zerstörungspotenzial für Mensch, Technik und Umwelt ausgeht (Perrow, 1999, nach Badke-Schaub et al., 2012, S. 25 f.).

Die Sicherheit eines soziotechnischen Systems ist nicht alleine von der technischen Zuverlässigkeit der einzelnen Komponenten abhängig, sondern vom Zusammenwirken technischer, organisatorischer und insbesondere auch menschlicher Faktoren (Gherardi, Nicolini & Odella, 1998, nach Badke-Schaub et al., 2012, S. 22). Diese Erkenntnis hat sich insbesondere in der Nuklearindustrie durchgesetzt und spiegelt sich beispielsweise in den Richtlinien der Internationalen Atomenergie-Organisation wider (IAEA, 2016, S. 19). Daraus resultierend werden in der Nuklearindustrie seit den 1980er-Jahren Methoden der Human Reliability Analyse (HRA) entwickelt und genutzt, mithilfe derer das Verhalten von Menschen in soziotechnischen Systemen modelliert und mögliche Fehler sowie deren Auswirkungen identifiziert und bewertet werden. Diese Notwendigkeit der Berücksichtigung des Menschen ist nicht auf die Nuklearindustrie begrenzt und wird auch in anderen Systemen vermehrt mitbedacht. Auch wenn HRA-Methoden schon früher außerhalb der Nuklearindustrie angewendet wurden, zeigen Patriarca, Ramos, et al. (2020, S. 11) in ihrem bibliometrischen Review, dass in wissenschaftlichen Publikationen erst seit 2009 Schlüsselwörter verwendet werden, die auf eine Anwendung in Bereichen wie z. B. Prozessindustrie schließen lassen.

Die Abteilung Risk Management Advisory der DNV Energy Systems Germany GmbH bietet Ihren Kunden Risikobewertungen an, z. B. mithilfe von qualitativen (Hazard Identification Analysen - HAZID / Hazard and Operability Analysen - HAZOP) und semi-quantitativen (Layers of Protection Analysis - LOPA) Methoden sowie Auswirkungsmodellierungen. Zu den Kunden gehören beispielsweise Anlagenbetreiber aus der Öl- und Gasindustrie sowie Betreiber von Anlagen zur Energieerzeugung aus erneuerbaren Energiequellen. Die bisherigen Dienstleistungen werden aufgrund der oben beschriebenen Relevanz des Human Factors durch HRA-Methoden ergänzt. Begründet durch die Erfahrungen des DNV entstand die Annahme, dass HRA-Methoden in der Öl- und Gasindustrie zwar angewendet werden, es allerdings noch keine einheitlichen Regularien, Guidelines oder Best Practices gibt. Somit bleibt die Frage nach der Übertragbarkeit dieser Methoden auf unterschiedliche Kontexte und Arbeitsumgebungen offen. Die Entwickler der Petro-HRA-Methode, die im Jahr 2017 veröffentlicht wurde, gehen beispielsweise davon aus, dass eine Anpassung für die Petroleumindustrie erforderlich ist, und schreiben:

HRA provides a systematic means of evaluating the human contribution to risk. Most of the methods that are used today were originally developed for the nuclear

industry, and this creates challenges when attempting to use such methods for analysis of major accident scenarios in petroleum. It is true that there are many similarities between the nuclear and petroleum industries; both are high-hazard industries, with sophisticated defence-in-depth risk management strategies and potentially devastating consequences if something does go wrong. But there are also significant differences between the two industries, especially when considering human actions in major accident scenarios. (Taylor et al., 2020, S. 2)

Diese Arbeit möchte diese Aussage durch eine strukturierte Literaturrecherche überprüfen und möglichst weitere Belege und Quellen finden. Auf der Literatur aufbauend sollen die vorgenommenen Anpassungen herausgearbeitet werden. Die Forschungsfrage lautet daher:

Wie werden HRA-Methoden für die Anwendung in der Öl- und Gasindustrie angepasst?

Perspektivisch besteht die Hoffnung, aus dem Prozess der Adaption von der Nuklearindustrie auf die Öl- und Gasindustrie lernen zu können, um die HRA auch im Kontext von anderen Systemen anwenden zu können.

Zur Beantwortung der Forschungsfrage werden im theoretischen Teil zunächst die grundlegenden Konzepte zur Beschreibung des menschlichen Einflusses auf die Systemsicherheit erläutert sowie die Human Reliability Analyse als Methode vorgestellt. Dabei wird auch auf die regulatorischen Rahmenbedingungen eingegangen. Anschließend wird das methodische Vorgehen der Literaturrecherche beschrieben. In Kapitel 4 werden die im Hinblick auf die Forschungsfrage relevanten Ergebnisse der Literaturrecherche präsentiert. Abschließend erfolgt in der Diskussion eine Gegenüberstellung und Bewertung dieser Ergebnisse.

2. Theoretischer Hintergrund

2.1 Sicherheit und der Faktor Mensch

In den Ingenieurwissenschaften ist Sicherheit meist als Freiheit von nicht vertretbaren Risiken definiert. Risiko wiederum ist als Kombination aus der Wahrscheinlichkeit des Schadensauftretens und der Schwere des Schadens definiert. Wenn ein System die vorgesehene Funktion ausfallfrei erfüllt, die Wahrscheinlichkeit für einen Zwischenfall oder Schaden also gering ist, kann es als zuverlässig beschrieben werden. Sicherheit wird daher auch oft mit Zuverlässigkeit gleichgesetzt (Badke-Schaub et al., 2012, S. 22). Für die Risikoanalyse technischer Systeme existieren verschiedene qualitative und quantitative Methoden. So werden bspw. im Rahmen von qualitativen bzw. semi-qualitativen HAZID-, HAZOP- oder LOPA-Analysen technische Ursachen für Gefahren identifiziert, deren Eintrittshäufigkeit quantifiziert und das Schadensausmaß qualitativ beschrieben. Im Rahmen von quantitativen Methoden wird zusätzlich das Schadensausmaß, z. B. die Ausbreitungszonen von Gasen, modelliert.

Für soziotechnische Systeme oder Organisationen reicht eine Betrachtung der technischen Zuverlässigkeit nicht aus. In diesen ergibt sich Sicherheit erst aus der erfolgreichen Interaktion verschiedenster Faktoren und insbesondere der Interaktion von Mensch und Technik. Dies scheint durch die gängigen Unfallstatistiken bestätigt zu werden, die besagen, dass der Mensch in der Mehrzahl der Fälle Ursache für Unfälle ist. Dabei muss allerdings berücksichtigt werden, dass sich letztlich alle Unfälle und deren zugrunde liegenden Fehler auf den Menschen zurückführen lassen (Badke-Schaub et al., 2012, S. 5). Der britische Chemiker Trevor Kletz, Pionier der Prozesssicherheit, beschrieb dies wie folgt: "To say accidents are due to human failing is like saying that falls are due to gravity. It is true but does not help us prevent them" (Kletz, 2001, S. 230, nach Selvik & Bellamy, 2020, S. 9). Für eine bessere Aussagekraft ist daher eine ausführlichere und differenziertere Betrachtung nötig.

Gesine Hofinger identifiziert als Kern aller Fehlerdefinitionen die Beschreibung als „eine Abweichung von einem als richtig angesehenen Verhalten oder von einem gewünschten Handlungsziel, das der Handelnde eigentlich hätte ausführen bzw. erreichen können.“ (Badke-Schaub et al., 2012, S. 40).

Die Sicherheit und Zuverlässigkeit von soziotechnischen Systemen kann also erst korrekt eingeschätzt und verbessert werden, wenn auch die möglichen menschlichen Fehler und deren Häufigkeit im Rahmen von Risikoanalysen berücksichtigt werden. Genau für diesen Zweck wurde die HRA entwickelt. Eine der elementaren Aufgaben einer HRA ist es, die möglichen Abweichungen im menschlichen Handeln zu identifizieren. Für die effektive Durchführung dieses Schritts der Human Error Identification (HEI), welcher in Kapitel 2.3 genauer erläutert wird, existieren eine Vielzahl von Taxonomien, die Fehler auf verschiedenen Ebenen betrachten. So kann bspw. zwischen der beobachtbaren Fehlerform (Was ist passiert?) und der auf theoretischen Annahmen beruhenden Fehlerart (Warum ist etwas passiert?) unterschieden werden (Badke-Schaub et al., 2012, S. 48). Das Ziel dieser Taxonomien ist es, eine Vielzahl beobachtbarer Fehler auf wenige Grundmuster zu reduzieren (Badke-Schaub et al., 2012, S. 40).

Die einfachste Unterteilung von Fehlerarten unterscheidet in Errors of Omission (EOO) und Errors of Commission (EOC). Diese wurde von Alan Swain und H. E. Guttman, den Entwicklern der *Technique for human error-rate prediction* (THERP) vorgeschlagen. EOOs sind das Unterlassen einer notwendigen Handlung, während EOCs die falsche Ausführung einer Handlung beschreiben (Liu et al., 2021, S. 529).

Als besonders einflussreich hat sich die Klassifizierung nach James Reason erwiesen, welche die Ebenen der Fehlerform und Fehlerart verbindet, indem er die vorangegangenen Arbeiten von Norman und Rasmussen kombiniert. Mithilfe dieser lassen sich, wie in Abbildung 1 dargestellt, Fehler in Slips (Aufmerksamkeitsfehler), Lapses (Gedächtnisfehler), Mistakes (regelbasierte Fehler) und Violations (Verstöße) unterteilen (Badke-Schaub et al., 2012, S. 52).

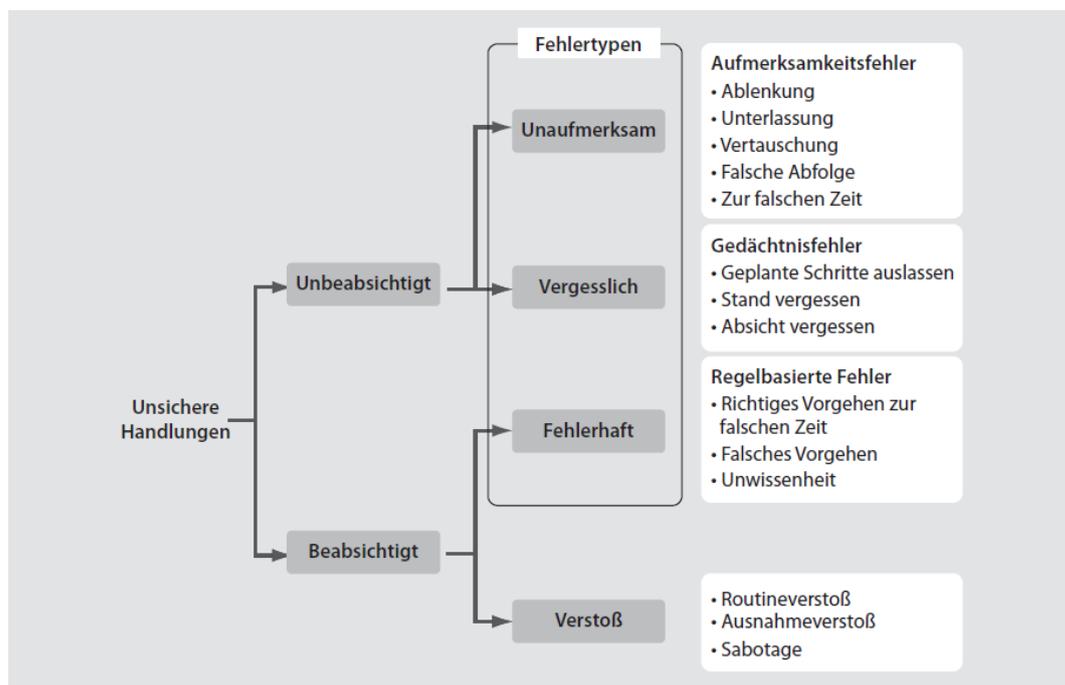


Abbildung 1: Fehlerklassifikation nach Reason
(Reason, 1990, nach Badke-Schaub et al., 2012, S. 53)

Slips und Lapses beschreiben Fehler in der Ausführung einer Handlung. Slips sind in diesem Modell Handlungen, die nicht korrekt ausgeführt werden, während Lapses unterlassene Handlungen beschreiben und als Ursache daher ein Gedächtnisfehler angenommen wird. Als Mistakes werden Planungsfehler beschrieben. Dabei sind geplante Handlungen bspw. aufgrund von fehlendem Wissen nicht dem Problem angemessen. Fehler werden nie vollständig verhindert werden können, Slips und Lapses können aber durch unterstützende Arbeitsbedingungen und Systemgestaltung verringert werden. Dies können regelmäßige Pausen zur Verbesserung der Aufmerksamkeit und technische Redundanz sein. Da Mistakes stark von kognitiven Prozessen abhängen, sind diese deutlich schwerer vorherzusagen und lassen sich z. B. durch Training und soziale Redundanz auffangen (Badke-Schaub et al., 2012, S. 50–54).

Zusätzlich berücksichtigt Reason mit der Klasse Violations (Verstöße) Handlungen, mit denen ein Mensch bewusst Regeln übergeht. Dies geschieht nicht mit dem Ziel der

Sabotage, sondern wird durch den Konflikt zwischen verschiedenen Zielen verursacht. Ein solcher Konflikt kann z. B. zwischen Sicherheit und wirtschaftlichem Druck bestehen, aber auch persönliche Bedürfnisse, ausgelöst durch Schlafmangel oder Statussicherung, betreffen (Reason, 1990, nach Badke-Schaub et al., 2012, S. 53 f.). Für gewohnheitsmäßige und akzeptierte Regelverstöße in einer Organisation hat Diane Vaughan den Begriff der *Normalization of deviance* (Routineverstöße) eingeführt (Vaughan, 1997, nach Badke-Schaub et al., 2012, S. 54).

Des Weiteren ist es sinnvoll, Fehler hinsichtlich ihrer räumlichen und zeitlichen Lokalisation innerhalb eines soziotechnischen Systems zu klassifizieren. Geschehen Fehler direkt an der Mensch-System-Schnittstelle spricht man auch vom *scharfen Ende* (*sharp end*) einer Organisation. Diese aktiven Fehler haben das Potenzial, Zwischenfälle direkt auszulösen, können aber auch einfacher entdeckt werden. Fehler, die räumlich und zeitlich von einem möglichen Unfall entfernt geschehen, d. h. am *stumpfen Ende* (*blunt end*), werden latente Bedingungen genannt. Unfälle werden mit der Kombination von aktiven Fehlern und latenten Bedingungen erklärt (Badke-Schaub et al., 2012, S. 43 f.). Zur Verdeutlichung dieses Umstandes bietet sich auch die Klassifikation von Meister an. Diese unterscheidet in vier Kategorien, in welchen Fehler ihren Ursprung haben können. *Operating Errors* geschehen am scharfen Ende, wenn ein System nicht nach Plan bedient wird. Weiterhin wird in *Design Errors*, *Manufacturing Errors* und *Installation and Maintenance Errors* unterschieden, welche zeitlich und räumlich von *Operating errors* getrennt sind, diese aber begünstigen oder latente Bedingungen schaffen (Meister, 1971, nach Badke-Schaub et al., 2012, S. 49).

In der Öffentlichkeit und Rechtsprechung werden aktive Fehler oft genutzt, um Unfälle zu erklären und verantwortliche Personen zu benennen und sanktionieren. Tatsächlich zeigen aber auch die beschriebenen Modelle einen starken Einfluss des Kontexts (regelmäßige Pausen, latente Bedingungen etc.) Diese Relevanz des Kontexts wird auch in der HRA anerkannt. Fehler werden nicht auf Inkompetenz, mangelnde Motivation oder Nachlässigkeit zurückgeführt, sondern auf verschiedene Faktoren des Systems, z. B. dem Design, der Umgebung, Kultur und Ausbildung. Versuche, die menschliche Zuverlässigkeit zu erhöhen, sollten sich daher auf die Verbesserung des Arbeitssystems konzentrieren. Der Mensch nimmt in diesem eine Rolle ein, in welcher Fehler meistens unabhängig vom Individuum sind (Kirwan, 2015, S. 819 f.).

In diesem Zusammenhang ist auch der Begriff Human Error bzw. menschlicher Fehler kritisch zu betrachten, da durch diesen die Relevanz des Kontexts nicht deutlich wird und damit eine Schuldzuweisung verbunden werden kann, die eigene Probleme für eine effektive Sicherheitskultur mit sich bringt. Dennoch wird dieser Begriff aus Mangel an passenden Alternativen in den meisten Publikationen und auch dieser Arbeit vorrangig genutzt (Kirwan, 2015, S. 793).

Es wird deutlich, dass zur Vermeidung von Fehlern und einer Erhöhung der Zuverlässigkeit vor allem eine Verbesserung der Kontextfaktoren nötig ist. Alternativ bzw. zusätzlich kann aber auch die Technik bspw. durch einen höheren Grad der Automatisierung verändert werden. Im Folgenden wird daher kurz beschreiben, wie solche Veränderung dennoch die Interaktion zwischen Menschen und Technik und damit die Zuverlässigkeit beeinflussen.

Wenn die Zuverlässigkeit von Systemen als Ergebnis der Interaktion von Hardware, Software und menschlichem Einfluss betrachtet wird, kann diese durch einen höheren Automatisierungsgrad zwar gesteigert werden, jedoch funktionieren Hardware und herkömmliche algorithmische Software nur innerhalb der für sie vorgesehenen Rahmenbedingungen. Wirken nicht bedachte Faktoren auf ein System oder fällt dieses in einen instabilen Zustand, werden Hard- und Software versagen und sich nicht selbstständig in einen stabilen Zustand (zurück)versetzen. Menschen sind hingegen in der Lage (eigene) Fehler zu erkennen und flexibel auf Probleme zu reagieren und diese eventuell zu lösen. Der Mensch kann also auch als Problemlöser oder Barriere gesehen werden (Liu et al., 2021, S. 516 f.).

In Bezug auf die Automatisierung von Systemen ist es auch sinnvoll die 1983 von Lisanne Bainbridge formulierten *Ironien der Automatisierung* zu berücksichtigen. Darin wird unter anderem darauf hingewiesen, dass mittels Automatisierung Aufgaben zuverlässiger durchgeführt werden sollen, diese aber weiterhin von Menschen überwacht werden und dieser in kritischen Situationen eingreifen soll. Bainbridge weist darauf hin, dass das Überwachen von Systemen teilweise höhere Anforderungen an die Menschen stellt und mehr Training benötigt. Aus einem höheren Grad an Automatisierung kann auch folgen, dass das Wissen des Menschen über ein System mit dem Grad an Automatisierung abnimmt, da sich der Mensch weniger intensiv mit dem System auseinandersetzen muss. Sobald das System dann aber in einen kritischen Zustand fällt, soll der Mensch korrigierend eingreifen (Bainbridge, 1983, S. 775 f.).

2.2 Abgrenzung HRA und Human Factor

Die bisher beschriebenen Konzepte werden typischerweise der Disziplin Human Factors (HF) zugeschrieben. Diese untersucht, basierend auf Erkenntnissen und Methoden aus verschiedenen Bereichen, wie zum Beispiel der Psychologie, die Interaktion zwischen Menschen und Technik. Dementsprechend könnte man vermuten, dass die HRA einen engen Bezug zur HF-Forschung und Anwendung hat. In der Praxis werden HRA-Methoden dennoch nicht der HF-Disziplin zugeordnet, sondern eher mit der Zuverlässigkeitstechnik (reliability engineering) verbunden. Dies zeigt sich beispielsweise darin, dass viele der Fortschritte im Bereich HRA außerhalb der üblichen Journals für Human Factors veröffentlicht werden. Es ist wahrscheinlicher, dass eine Arbeit im Bereich HRA in der Zeitschrift *Reliability Engineering and System Safety* erscheint, als in der Zeitschrift *Human Factors* (Boring, 2012, Kapitel 3.3; Patriarca, Ramos, et al., 2020, S. 5).

Historisch betrachtet entstand die HRA im Rahmen von Bemühungen, die Zuverlässigkeit bei der Montage von Atomwaffen in den 1950er-Jahren zu erhöhen. HF wird hingegen auf Bemühungen der Verbesserung der menschlichen Leistung zugeschrieben, insbesondere der Entwicklung besserer Landkontrollen für Kampfflugzeuge während des Zweiten Weltkriegs. Diese historischen Unterschiede in den Ursprüngen haben zu unterschiedlichen Forschungsansätzen geführt. Während HF in der Regel eine starke Basis in der empirischen Beobachtung hat, um Empfehlungen für die Gestaltung von Arbeitssystemen zu generieren, stützt sich HRA in der Regel stark auf Prozesswissen und Expertenschätzungen (Boring, 2012, Kapitel

3.3). Letzteres wurde aber auch früh kritisiert und es wurden Versuche unternommen, HRA auf Basis von empirischen Daten durchzuführen.

2.3 Allgemeiner Prozess der HRA

Für die generelle Beschreibung einer HRA beziehen sich viele Publikationen auf Barry Kirwan (Kirwan, 2015, S. 795–800), welcher einen allgemeinen HRA-Prozess formuliert hat. Dieser besteht aus 10 Schritten und wird in Abbildung 2 verdeutlicht.

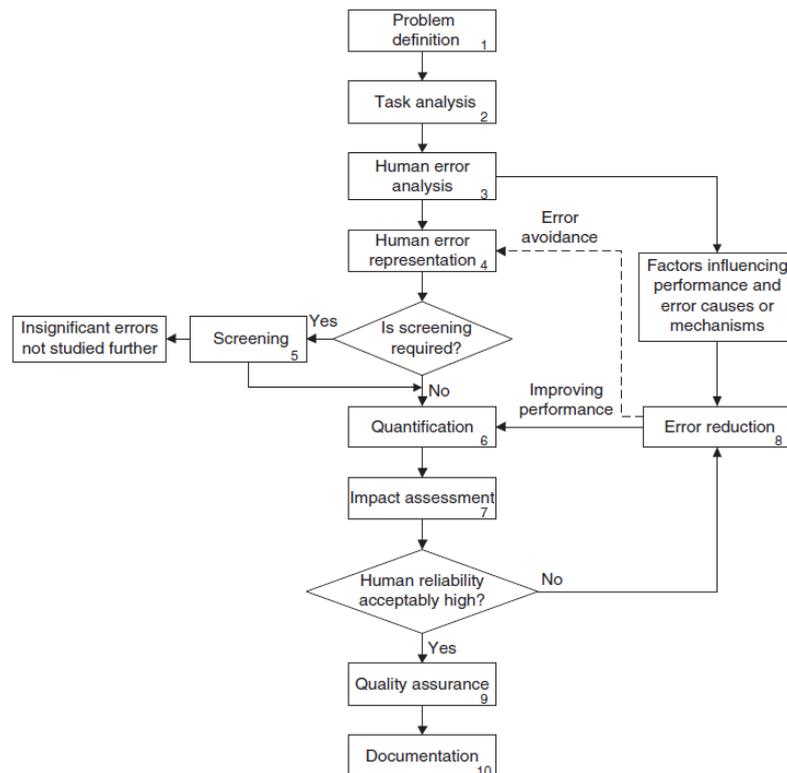


Abbildung 2: Schematische Darstellung des HRA-Prozesses (Kirwan, 1994, nach Liu et al., 2021)

Im Rahmen der *Problem Definition* soll der Umfang der HRA geplant und festgelegt werden. Relevante Fragen sind hier unter anderem: *Wird die HRA als Teil eines QRAs oder eigenständig (stand alone) durchgeführt?* oder *Welche Ressourcen sind für die Durchführung verfügbar?*

Mithilfe der *Task Analysis* sollen die korrekten und erforderlichen Handlungen der mit dem System interagierenden Menschen beschrieben werden. Für die Analyse und Darstellung können verschiedene Formate wie z. B. die Hierarchical Task Analysis, Tabular Task Analysis, Timeline Analysis oder Link Analysis genutzt werden. Mithilfe dieser Formate werden ein Ziel bzw. eine Aufgabe in Unterziele & -aufgaben zerlegt, bis die zur Aufgabenerfüllung notwendigen mentalen und physischen Prozesse erkennbar werden.

Auf Basis der Task Analysis werden im Rahmen der *Human Error Analysis* (Schritt 3 in Abbildung 2) mögliche Human Failure Events (HFE) identifiziert. Weitere Bezeichnung für diesen Schritt lauten *Human Error Identification (HEI)* oder *Human Error Prediction*. Für bestehende Systeme kann und sollte die Fehleridentifikation von bestehenden Erfahrungen und vergangenen Zwischenfälle ausgehen. Dafür könne

Interviews geführt und ausgewertet werden, oder falls vorhanden, Dokumentation, Unfallberichte oder Informationen aus einem Meldesystem über Beinahe-Ereignisse und Zwischenfälle ausgewertet werden. Im Rahmen von Interviews gilt es Verzerrungen und Beeinflussung zu erkennen bzw. zu verhindern (Kirwan, 2015, S. 800). Es existieren wiederum eigenständige Methoden, welche die HEI strukturieren und in die HRA eingebettet werden können. Beispiele dafür sind SHERPA und PHEA, zwei verwandte Techniken oder TRACER. Ein Auszug möglicher Fehlertypen der PHEA-Methode ist in Tabelle 1 abgebildet, in welcher die Klassifikation auf der Ebene der Fehlerform stattfindet.

| Error types | Error code | Explanation |
|-----------------|------------|---|
| Planning errors | P1 | Ignore preconditions of a plan |
| | P2 | Execute wrong plan |
| | P3 | Execute inappropriate plan |
| | P4 | Right plan but conducted in wrong order |
| | P5 | Execute correct plan too soon/late |
| Action errors | A1 | Action too late |
| | A2 | Action too early |
| | A3 | Action omitted |
| | A4 | Action too fast |
| | A5 | Action too slow |
| | A6 | Action too much |
| | A7 | Action too little |
| | A8 | Action incomplete |
| | A9 | Action in wrong direction |
| | A10 | Wrong action on correct object |

Tabelle 1: Auszug aus der PHEA Error Taxonomy (Liu et al., 2021, p. 531)

Mithilfe der *Human Error Representation* werden die HFE in den Kontext des Systems gesetzt und die Wirkungen auf das System dargestellt. Dafür werden bspw. Ereignisbaumanalysen oder Fehlerbaumanalysen genutzt, welche eine Berechnung von Wahrscheinlichkeiten ermöglichen. Diesen HFEs können im *Screening* testweise pessimistische Human Error Probabilities (HEP), z. B. $HEP = 1$, zugeordnet werden. Aufgaben, die auch dann keine sicherheitsrelevanten Auswirkungen haben, können von der Analyse ausgeschlossen werden.

Auf die bisher beschriebenen qualitativen Schritte folgt die Zuordnung von HEPs, d. h. Schritt 6 des in Abbildung 2 gezeigten HRA Prozesses. Theoretisch sollten HEPs auf der beobachteten relativen Häufigkeit von HFEs basieren:

$$HEP = \frac{\text{number of errors observed}}{\text{number of opportunities for error}}$$

Diese frequentistische Wahrscheinlichkeitsdefinition, die auf dem Gesetz der großen Zahlen beruht, setzt theoretisch eine unendliche Anzahl an Beobachtungen voraus. In der Praxis ist es jedoch nicht realistisch, den Zähler und Nenner für jedes System individuell zu bestimmen. Daher ist die Ermittlung der HEP auf Grundlage generalisierter Daten oder Schätzungen ein wesentlicher Bestandteil der HRA.

Da sich, wie bereits beschrieben, die Erkenntnis durchgesetzt hat, dass der Kontext die Kognition und das Verhalten beeinflussen, werden auch im Rahmen der HRA Faktoren genutzt, welche die HEP beeinflussen. Diese werden je nach Methode bspw.

Performance Influencing Factors (PIF) in der *Standardized Plant Analysis Risk-HRA* (SPAR-H) und Petro-HRA, Performance-Shaping Factors (PSF) in THERP oder Error Producing Conditions (EPC) in der *Human Error Assessment and Reduction Technique* (HEART) genannt.

Nach der Quantifizierung wird der Einfluss der HEPs auf das Gesamtrisiko betrachtet und evtl. mit einem vorher festgelegten akzeptiertem Grenzkrisiko verglichen. Gegebenenfalls werden Ansätze zur Risikoreduzierung gesucht.

Der hier beschriebene Ablauf ist ein idealisierter Prozess. Abhängig vom Kontext oder der verwendeten Methode werden evtl. nur einzelne Schritte durchgeführt.

2.4 Historische Entwicklung und Klassifizierung

Die historische Entwicklung von HRA-Methoden lässt sich, wie einleitend beschrieben, bis in die 1950er zurückverfolgen. Im Kontext der Zuverlässigkeit bei der Montage von Atomwaffen beschäftigte sich Swain mit der Prognose von menschlichen Fehlern. Im Rahmen der Jahresversammlung der Human Factors Society wurde im November 1962 eine frühe Version von THERP vorgestellt. Die Idee, diese Methode auch in der Kernenergie anzuwenden, gipfelte 1983 in der Veröffentlichung *NUREG/CR 1278 - Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications* der *U.S. Nuclear Regulatory Commission* (Boring, 2012, Kapitel 2). Seitdem wurden mehr als 50 HRA-Methoden veröffentlicht (Bell und Holroyd, 2009, nach Liu et al., 2021, S. 542). Die Klassifizierung dieser Methoden kann anhand verschiedener Kriterien erfolgen. Häufig zu finden ist eine Unterteilung in Generationen, welche auf chronologischen und inhaltlichen Kategorien beruht, in der Einteilung aber nicht konsistent ist. Methoden der 1. Generation sind von der technischen Zuverlässigkeitsanalyse geprägt und betrachten decken daher insbesondere EOOs auf. Dazu zählen THERP und HEART ((Liu et al., 2021, S. 542 f.)). Methoden der 2. Generation wird eine Berücksichtigung der Kognition, des Kontexts und von EOCs zugeschrieben. Diese Entwicklung ging auch mit der kognitiven Wende in der Psychologie, also dem Paradigmenwechsel weg vom Behaviorismus hin zum Kognitivismus, einher (Boring, 2012, Kapitel 3.2; Liu et al., 2021, S. 543). Als Methoden der 2. Generation können *A technique for human error analysis* (ATHEANA), *Cognitive Reliability Error Analysis Method* (CREAM) von Erik Hollnagel und *MERMOS* genannt werden (Kirwan, 2015, S. 818). Ronald Boring schlägt Kriterien für eine dritte Generation der HRA-Methoden vor. Während die Methoden der ersten und zweiten Generation einen statischen Zeitpunkt betrachten, sollen in den Methoden der dritten Generation die PSFs und damit die HEP als Funktionen der Zeit dargestellt werden. Diese werden daher *dynamische HRA* genannt (Boring, 2007, Kapitel 3). Eine Suche auf Scopus am 29.08.2014 mit dem Suchstring "dynamic HRA" ergab allerdings nur 30 Publikationen seit 2006.

Die verschiedenen Methoden unterscheiden sich vor allem in der Aufgaben-/Fehlerklassifizierung, den Kontextfaktoren und der Berechnung der HEPs. Exemplarisch wird hier die Quantifizierung von SPAR-H und HEART vorgestellt. Während die SPAR-H die analysierte Handlung in *Diagnosis*, *Action* oder einer Kombination aus beidem unterteilt, werden Handlungen in HEART einer von neun in Tabelle 2 abgebildeten *Generic Task Types* (GTT) zugeordnet. Diesen Kategorien sind

wiederum Fehlerwahrscheinlichkeiten zugeordnet. In SPAR-H betragen diese Nominal HEP (NHEP) genannten Wahrscheinlichkeiten für *diagnosis tasks* 0.01 und 0.001 für *action tasks*. Im Rahmen von HEART werden diese *nominal human unreliability* oder GT_{HEP} genannt (Liu et al., 2021, S. 545–548).

Mithilfe von acht PSFs in SPAR-H bzw. 38 EPCs in HEART wird der Einfluss von weiteren Einfluss- und Kontextfaktoren berücksichtigt. Diese werden von den Analysten gewichtet. In SPAR-H sind für jeden PSF mehrere Level beschrieben, denen jeweils ein Faktor (Multiplier) zugeordnet ist. In HEART ist jedem EPC ein maximaler Effekt zugeordnet, dessen Gewichtung über einen zusätzlichen Faktor erfolgt. Dieser *Assessed Proportion Of Affect (APOA)* kann zwischen 0,1 (sehr schwacher Effekt) und 1 (voller Effekt) liegen und muss von den Analysten eingeschätzt werden. Aus diesem unterschiedlichen Vorgehen ergeben sich verschiedene Formeln für die Berechnung der HEP. Für HEART lautet die Formel:

$$HEP = GT_{HEP} \times \prod WF_i$$

mit

$$WF = (EPC_i - 1) \times APOA_i + 1$$

In SPAR-H ergibt sich die HEP aus der Multiplikation der NHEP und den Faktoren der PSFs (Liu et al., 2021, S. 545–548).

Für eine weitere Übersicht der existierenden Methoden lässt sich auch eine Einordnung in deren Herkunft bzw. Finanzierungsquellen vornehmen. Als wichtige Organisationen sind die in Abbildung 3 aufgeführte *U.S. Nuclear Regulatory Commission* (U.S. NRC) und das ebenfalls amerikanische, unabhängige *Electric Power Research Institute* (EPRI) zu nennen (Boring, 2012, Kapitel 3). Aus der Abbildung wird auch die Relevanz amerikanischer Beiträge für den Themenbereich ersichtlich. Dies wird auch anhand der Anzahl an wissenschaftlichen Publikationen aus dem amerikanischen Raum deutlich. Ramos et al. (2022, S. 913) zeigen, dass von den bei Scopus unter dem Suchstring *“human reliability” OR “human unreliability”* bis Dezember 2020 indexierten Veröffentlichungen (n=2487) 22,29 % aus den USA stammen, seit 2000 aber auch die Anzahl an Publikationen aus Asien zunimmt. So wurden 11,36 % der Publikationen chinesischen Autoren und 5,06 % koreanischen Autoren zugeordnet.

Auch die Relevanz der Nuklearindustrie wird in Abbildung 3 sichtbar und kann anhand von bibliometrischen Daten belegt werden. Patriarca, Ramos, et al. (2020, S. 11) zeigen, dass andere Anwendungsfelder, wie bspw. die Prozessindustrie in der wissenschaftlichen Literatur erst ab den späten 2000er-Jahren vermehrt erwähnt werden. Allerdings gilt die Einschränkung, dass anhand von wissenschaftlicher Literatur nicht zwingend die tatsächliche Anwendung durch die Industrie abgeleitet werden kann (Patriarca, Ramos, et al., 2020, S. 2).

| Generic task type | Proposed nominal human unreliability (5th–95th percentile bounds) |
|--|---|
| Totally unfamiliar, performed at speed with no real idea of likely consequences | 0.55 (0.35–0.97) |
| Shift or restore system to a new or original state on a single attempt without supervision or procedures | 0.26 (0.14–0.42) |
| Complex task requiring high level of comprehension and skill | 0.16 (0.12–0.28) |
| Fairly simple task performed rapidly or given scant attention | 0.09 (0.06–0.13) |
| Routine, highly-practiced, rapid task involving relatively low level of skill | 0.02 (0.007–0.045) |
| Restore or shift a system to original or new state following procedures, with some checking | 0.003 (0.0008–0.007) |
| Completely familiar, well-designed, highly practiced, routine task occurring several times per hour, performed to highest possible standards by highly-motivated, highly-trained and experienced person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant job aids | 0.0004 (0.00008–0.009) |
| Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system stage | 0.00002 (0.000006–0.009) |
| Miscellaneous task for which no description can be found | 0.03 (0.008–0.11) |

Tabelle 2: Generic Task Types in HEART (Williams, 1988, nach Liu et al., 2021, S. 546)

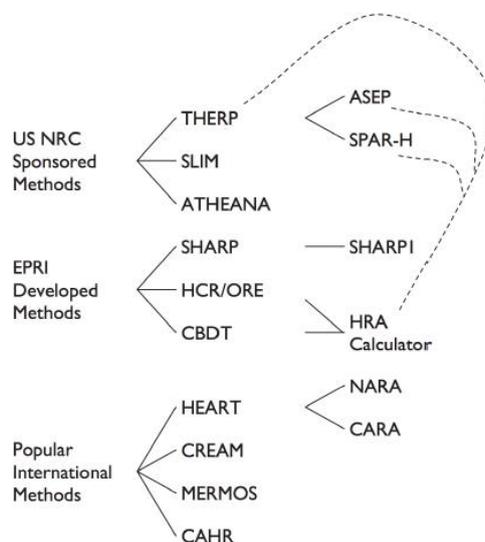


Abbildung 3: HRA-Methoden nach Finanzierungsquellen (Boring, 2012, Kapitel 3)

2.5 Trends und Herausforderungen

Als zukünftige Trends identifizieren Patriarca, Ramos, et al. (2020, S. 14) insbesondere die empirische Erhebung von HEPs und dem Einfluss von PSFs, den Umgang mit Unsicherheit dieser Daten und die Nutzung von Bayes'sches Netzen (BN).

Auch Hou et al. (2021, S. 10 f.) identifiziert zukünftige Entwicklungen und Forschungsschwerpunkte:

1. Die Erhebung und Nutzung empirischer Daten
2. Der Umgang mit Unsicherheit und der Abhängigkeit von Expertenmeinung
3. Standardisierung von PSFs und Abhängigkeiten zwischen PSF
4. Anpassung von HRA-Methoden an digitale Umgebungen
5. Der ausführliche Vergleich der verschiedenen Methoden
6. Die Nutzung von künstlicher Intelligenz und neuronalen Netzen

Anhand dieser Schwerpunkte lassen sich auch die aktuellen Herausforderungen erkennen. Im Folgenden wird ausführlicher auf BNs und die empirische Fundierung der HRA eingegangen.

Ein BN ist eine Struktur, mit der sich bedingte Wahrscheinlichkeiten, d. h. Abhängigkeiten zwischen mehreren Variablen, beschreiben lassen. Die Variablen werden durch sogenannte Knoten, die Abhängigkeiten zwischen den Variablen werden durch Kanten repräsentiert (Mkrtchyan et al., 2015, S. 3). Die meisten HRA-Methoden treffen die vereinfachte Annahme, dass die PSFs voneinander unabhängig sind. In der Realität wird aber bspw. der Einfluss der Variable *Task Complexity* auf die HEP durch die Variablen *Procedure Guidance*, *Experience* oder *Time availability* moderiert. Die PSFs sind also nicht unabhängig voneinander (Mkrtchyan et al., 2015, S. 6; Liu und Li, 2014, nach Liu et al., 2021, S. 555). BN können auch helfen, komplexe Abhängigkeiten zwischen HFES zu modellieren, wie z. B. den Einfluss eines Fehlers in einer Aufgabe A auf eine andere Aufgabe B. Die Struktur des BNs kann auf Basis von Daten oder der Einschätzung von Experten entwickelt werden (Mkrtchyan et al., 2015, S. 2 f.). Vorteile von BN sind unter anderem, dass sie komplexe Abhängigkeiten modellieren und gleichzeitig grafisch intuitiv darstellen können, verschiedene Datenquellen in ein Modell integrieren und Daten aktualisieren sowie vorwärts- und rückwärtsgerichtete Schlüsse ziehen können. (Mkrtchyan et al., 2015, S. 12; Golestani et al., 2020, S. 9).

Bemühungen, die HRA durch den Einsatz empirischer HEPs zu verbessern, zielen darauf ab, die ökologische Validität zu erhöhen. Daten können bspw. auf Basis von Vorfallberichten oder Simulationen erhoben werden. Allerdings haben sich die Versuche, Daten auf Basis von Vorfallberichten zu sammeln, als wenig praktikabel erwiesen. Dagegen gibt es mit der SACADA- und HuREX-Datenbank andauernde Projekte der U.S. NRC und des Korean Atomic Energy Research Institute zur Erhebung simulationsbasierter Daten (Liu et al., 2021, S. 556 f.).

Der Fokus auf die empirische Fundierung wird aber auch kritisch diskutiert. Laumann et al. (2018) thematisieren dazu die Frage, was HRA-Methoden insgesamt leisten können und sollen. Damit wird auch die Frage nach der Klassifizierung in qualitative oder quantitative Methoden thematisiert. Kritisiert werden einerseits die Ebene der Anwendung von HRA-Methoden und andererseits die Ebene der Erhebung von zu Grunde liegenden Daten der menschlichen Zuverlässigkeit und deren Darstellung in Datenbanken. Bezüglich der Anwendung können HRA-Methoden nicht die Gütekriterien quantitativer psychologischer Tests erfüllen, da aus einer mangelnden Standardisierung (Objektivität) eine mangelnde Reliabilität und Validität resultiert. Am Beispiel von SPAR-H und HEART wird weiter gezeigt, dass die inhaltlichen Konstrukte, die gemessen werden sollen, wie beispielweise der PSF *Available Time* nicht ausreichend definiert und damit nicht operationalisierbar sind. Dies führt vor allem zu einer mangelnden inhaltlichen Validität und einer mangelnden Interrater-Reliabilität, wodurch auch eine Überprüfung der Konstrukt- und Kriteriumsvalidität verhindert wird (Laumann et al., 2018, S. 316 f.).

In Bezug auf Bemühungen, empirische Daten der menschlichen Zuverlässigkeit zu erheben, wird kritisiert, dass HRA-Methoden als Antwort auf fehlende empirische Daten entwickelt wurden, nun aber ihrerseits versuchen, diese Daten empirisch zu

erheben (Laumann et al., 2018, S. 315 f.). Herausfordernd in experimentellen (simulatorbasierten) Erhebungen ist hier auch wieder die mangelnde Definition der Konstrukte, welche das Design von Experimenten erschwert. Weiterhin ist es z. B. sehr schwer, PSFs als unabhängige Variable zu kontrollieren und manipulieren. Auch ist der Ressourcenaufwand zum Nachweis einer Fehlerwahrscheinlichkeit von bspw. 10^{-3} sehr groß (Boring et al., 2009, S. 613).

Für die Darstellung von empirischen Daten in einer Datenbank muss die Struktur dieser auf Grundlage von bestimmten Konstrukten und deren Definitionen entwickelt werden. Da sich aber bspw. die Definitionen von PSF zwischen verschiedenen HRA-Methoden unterscheiden, können Datenbanken nur methodenspezifisch genutzt werden, was wieder mit einem hohen Ressourcenaufwand verbunden ist (Laumann et al., 2018, S. 318 f.).

Die in Kapitel 2.1 beschriebenen Unterschiede zwischen dem Ansatz der HF- und HRA-Forschung erschweren auch die Übertragung von Daten aus HF-Studien. Julius Persensky, ehemaliger Leiter der Human Factors-Forschung bei der U.S. NRC, verdeutlicht dies:

Human factors research should be used to provide data for HRA. Yet, there is tremendous cost associated with collecting original data or even reprocessing existing data for use in quantifying HRA. [...] HRA has not provided a clear prescription for human factors on what data are needed. What values are needed under what conditions? Moreover, HRA deals with extremely small probabilities of human actions, which can be extremely difficult to observe in human factors studies. To observe such low frequency human errors, it is necessary to run many more human participants in studies or to run them longer than we have in the past. Clear guidance on HRA data gathering requirements would facilitate research that is more attuned to helping HRA methods develop empirically driven HEPs. (Boring et al., 2009, S. 613)

Da sich die Erfüllung der quantitativen Gütekriterien also als sehr herausfordernd darstellen kann, plädieren Laumann et al. (2018, S. 320 f.) dafür, HRA als qualitative Methode wahrzunehmen. Eine ähnliche Bewertung nimmt Gould et al. (2012, S. 2018 f.) vor. Die qualitative Analyse, insbesondere die Aufgabenanalyse (HTA), ist besonders wichtig für die Identifikation vorher unbekannter Risiken. Die aus der quantitativen Analyse resultierenden HEPs sind aber eher relativ bzw. ordinal als absolut zu interpretieren und dienen damit vor allem der Priorisierung von risikoreduzierenden Maßnahmen.

2.6 Regulatorische Rahmenbedingungen und Standards

Auf europäischer Ebene dient die Richtlinie 2012/18/EU, auch Seveso-III-Richtlinie bezeichnet, als rechtliche Grundlage für die Beherrschung von Gefahren schwerer Unfälle mit gefährlichen Stoffen. In Deutschland erfolgte die Umsetzung dieser insbesondere durch Änderung des Bundes-Immissionsschutzgesetz (BImSchG) und der Störfall-Verordnung (12. BImSchV). Kriterium für die Anwendung der Störfall-Verordnung sind das Vorhandensein und Überschreiten festgelegter Schwellen von gefährlichen Stoffen (§ 1 12. BImSchV). Diese Kriterien treffen bspw. auf viele

Unternehmen der Prozessindustrie zu. Ausgeschlossen sind nach § 1 Abs. 3 12. BImSchV allerdings u. a.

- die Offshore-Erkundung und -Gewinnung von Mineralien, einschließlich Kohlenwasserstoffen sowie
- die unterirdische Offshore-Speicherung von Gas sowohl in eigenen Lagerstätten als auch an Stätten, wo auch Mineralien, einschließlich Kohlenwasserstoffe, erkundet und gewonnen werden.

Aus §8 und Anhang III Ziffer 2b der 12. BImSchV ergibt sich die Notwendigkeit für ein Sicherheitsmanagementsystem und die „Anwendung von Verfahren zur systematischen Ermittlung der Gefahren von Störfällen [...], sowie Abschätzung der Wahrscheinlichkeit und der Schwere solcher Störfälle“. Die Notwendigkeit der Betrachtung möglicher menschlicher Einflüsse wird nicht explizit erwähnt. Einzig in der nach § 19 12. BImSchV geforderten Meldung an die zuständigen Behörden nach Eintritt eines Ereignisses muss eine Ursachenklassifizierung in betriebsbedingte Ursachen, menschlicher Fehler, umgebungsbedingte Ursachen oder sonstige Ursachen vorgenommen werden.

Präzisiert werden das BImSchG und die dazugehörigen Verordnungen durch die Kommission für Anlagensicherheit (KAS), welche durch § 51a BImSchG gebildet wurde und dem Stand der Sicherheitstechnik entsprechende Regeln vorschlagen soll. Von dieser wird im *Leitfaden zum Konzept zur Verhinderung von Störfällen und zum Sicherheitsmanagementsystem* eine systematische Untersuchung aller technischen, organisatorischen und menschlichen Fehler empfohlen, die zu einem Notfall führen können (KAS, 2018, S. 37). Als beispielhafte Methode zur Ermittlung und Bewertung wird u. a. die HAZOP-Analyse genannt (KAS, 2018, S. 27 f.). Diese wird in der Praxis, wie in Kapitel 2.1 beschrieben, allerdings nicht für die Betrachtung menschlicher Einflüsse verwendet. Der Leitfaden *Kompetenzen bezüglich menschlicher Faktoren im Rahmen der Anlagensicherheit (Betreiber, Behörden und Sachverständige)* empfiehlt eine Kompetenzentwicklung in den vier Kompetenzfeldern Mensch, Organisation und Management, Gruppe sowie Arbeitsplatz und Arbeitsaufgabe (KAS, 2011, S. 6 ff.). Basierend auf Nr. 2a der Anlage III der 12. BImSchV wird Betreibern von Anlagen empfohlen, Gefahren von Störfällen unter Einbeziehung von Kompetenzen im Bereich menschlicher Faktoren zu ermitteln und bewerten (KAS, 2011, S. 22). Für beispielhafte Methoden wird auf den Leitfaden *Empfehlung für interne Berichtssysteme* verwiesen. Dieser spricht aber nur Empfehlungen für die retrospektive Analyse von Ereignissen aus. Methoden der prospektiven Analyse der menschlichen Zuverlässigkeit werden nicht thematisiert (KAS, 2008).

Für die nach § 1 Absatz 3 12. BImSchV ausgeschlossenen Offshore-Betriebe im Gebiet der Küstengewässer und des Festlandssockels der Bundesrepublik Deutschland gilt die Offshore-Bergverordnung (OffshoreBergV). Diese wurde zur Umsetzung der *Richtlinie 2013/30/EU des Europäischen Parlaments und des Rates vom 12. Juni 2013 über die Sicherheit von Offshore-Erdöl- und -Erdgasaktivitäten und zur Änderung der Richtlinie 2004/35/EG* erlassen. Gegenstand der Richtlinie ist die Festlegung von Mindestanforderungen für die Verhinderung schwerer Unfälle bei Offshore-Erdöl- und -Erdgasaktivitäten und die Begrenzung der Folgen derartiger Unfälle. In dieser wird vom Unternehmer ein systematisches Risikomanagements

sowie die Erstellung eines Berichts über ernste Gefahren gefordert (§40 Absatz 1, §43 OffshoreBergV). Dieser Bericht muss einen Nachweis darüber enthalten, dass alle ernstesten Gefahren sowie deren Eintrittswahrscheinlichkeit und Folgen ermittelt wurden (Anlage 1 Nr. 2 OffshoreBergV). Eine Identifikation und Betrachtung von menschlichen Fehlern wird nicht explizit gefordert, sollte aber durch die Anforderung der Betrachtung *aller ernstesten Gefahren* inkludiert sein.

Übergeordnet wird die Gewinnung von Kohlenwasserstoffen im Bereich des Festlandssockels in Deutschland durch das Bundesberggesetz (BBergG) und die Allgemeine Bundesbergverordnung geregelt. In Bezug auf die Analyse von Risiken wird der Unternehmer verpflichtet,

unter Beachtung der allgemein anerkannten sicherheitstechnischen, arbeitsmedizinischen und arbeitshygienischen Regeln sowie der sonstigen gesicherten arbeitswissenschaftlichen Erkenntnisse die erforderlichen Maßnahmen und Vorkehrungen zu treffen, um Beschäftigte und Dritte vor Gefahren für Leben, Gesundheit und Sachgüter zu schützen, soweit die Eigenart des Betriebes dies zulässt. (§61 Absatz 1 Satz 1a BbergG)

Insgesamt wird deutlich, dass die Betrachtung des menschlichen Einflusses auf die Sicherheit im deutschen Recht selten explizit erwähnt wird, aber als dem Stand der (Sicherheits)Technik entsprechend betrachtet werden kann.

Eine Betrachtung der nationalen Regulierung in Norwegen und dem Vereinigten Königreich, Länder, in denen die Förderung von Kohlenwasserstoffen eine deutlich größere wirtschaftliche Bedeutung hat, zeigt, dass dort die Berücksichtigung des menschlichen Faktors im Rahmen von *Quantitative Risk Assessments* (QRAs) zwar vorgeschrieben wird, die Art der Umsetzung aber offengelassen wird (Skogdalen & Vinnem, 2011, S. 472 f.). Dementsprechend fordert auch der relevante Standard der norwegischen Petroleumindustrie *NORSOK Z-013 Risk and emergency preparedness assessment* (2010, S. 21) eine mindestens qualitative Betrachtung der menschlichen und organisatorischen Faktoren. In der Praxis werden diese Anforderungen nur begrenzt umgesetzt. Eine von Skogdalen und Vinnem (2011, S. 476 f.) durchgeführte Analyse von 15 in Norwegen durchgeführten QRAs attestiert diesen eine größtenteils oberflächliche Betrachtung des HF. Auch äußern die Autoren die Vermutung, dass sich dieser Umstand erst durch Audits und Kontrollen durch die regulierenden Behörden ändern wird (Skogdalen & Vinnem, 2011, S. 477).

Auf zwischenstaatlicher bzw. internationaler Ebene ist die Betrachtung des menschlichen Einflusses auf die Sicherheit insbesondere wieder in der Nuklearindustrie festgelegt. Die *International Atomic Energy Agency* (IAEA), deren Sicherheitsstandards als Mindeststandards und Orientierungspunkte für die Regulierung auf nationaler Ebene dienen, hat in ihren *General Safety Requirements* die Notwendigkeit der Bewertung des menschlichen Faktors im Rahmen eines *Safety Assessments* festgelegt (IAEA, 2016, S. 19 f.). Als geeignete Methode wird die HRA explizit durch den *Specific Safety Guide-3* beschrieben (IAEA, 2024, S. 42 f.). Für die wenigsten Industrien existieren ähnlich institutionalisierte und multilaterale Richtlinien wie für die Nuklearindustrie. Für die Öl- und Gasindustrie wären internationale Abkommen vor allem in Bezug auf die Offshore-Industrie relevant, da von dieser im

Gegensatz zur landseitigen Förderung größere internationale und grenzüberschreitende Risiken ausgehen können (Rochette, 2014, S. 1). Ein Artikel über die rechtlichen Grundlagen eines französischen politischen Forschungsinstituts aus dem Jahr 2014 beschreibt eine Regelungslücke auf internationaler Ebene in Bezug auf die Sicherheit der Offshore-Industrie (Rochette, 2014, S. 2). Im Rahmen dieser Arbeit konnten keine Gesetze und Abkommen gefunden werden, die diese Lücke innerhalb der letzten Dekade geschlossen haben.

Ergänzend zur staatlichen Regulierung empfiehlt der *Report to the President der National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling* (2011, S. 234 f.) die Verbesserung der Sicherheits- und Risikomanagementstandards durch eine ausgeprägtere Selbstkontrolle (self-policy). Als Positivbeispiel wird die Nuklearindustrie genannt.

Auf Ebene der Standards wird die HRA in der *DIN EN IEC 31010:2022-09 – Risikomanagement - Verfahren zur Risikobeurteilung* genannt. Unterstützend für die Anwendung in der Öl- und Gasindustrie kann die Norm *ISO 14224:2016 - Petroleum, petrochemical and natural gas industries — Collection and exchange of reliability and maintenance data for equipment* betrachtet werden, welche Standards für die Datenerhebung nennt. Seit der aktuellen Version werden in dieser auch menschliche Fehler berücksichtigt. In einem Review der Norm durch Selvik und Bellamy (2020, S. 7 ff.) bewerten diese die bereitgestellten Informationen aber als nicht ausreichend für eine tatsächliche Analyse der menschlichen Handlungen.

3. Methode

Für die Literatursuche geeigneter Publikationen wurde auf die Datenbanken ScienceDirect, SpringerLink, Sage Journals und Scopus zurückgegriffen. Der Suchstring lautete:

"Human Reliability Analysis" AND (Oil OR Gas OR Petroleum).

Da die HRA eine sehr angewandte Methode ist, wurden keine Einschränkungen hinsichtlich des Dokumententyps vorgenommen, sodass insbesondere auch Konferenzberichte berücksichtigt wurden, die evtl. kein Peer-Review Verfahren durchlaufen haben. Inhaltliches Kriterium war, dass die Publikationen HRA-Methoden im Kontext der Öl- und Gasindustrie behandelten. Publikationen, welche die Verwendung der HRA in anderen Industrien wie der maritimen Industrie und damit auch der Tankerschifffahrt behandeln, wurden ausgeschlossen. Weiteres Ausschlusskriterium war die retrospektive Anwendung von HRA-Methoden, d. h. im Kontext von Vorfällen. Es wurden alle Publikationen innerhalb der letzten zehn Jahre bis zum 12.05.2024 berücksichtigt.

Der Ablauf der Literaturrecherche wird in der an das *PRISMA Flow Chart Diagram* angelehnten Abbildung 4 deutlich. Insgesamt wurden 793 Publikationen durch die vier genannten Datenbanken gefunden und 13 Dokumente manuell hinzugefügt. Die hohe Anzahl der über ScienceDirect gefundenen Publikationen resultiert daraus, dass die Suche über das Suchfeld *Find articles with these terms* durchgeführt wurde, wodurch alle Dokumententeile mit Ausnahme der Referenzen durchsucht wurden. Diese erweiterte Suchstrategie wurde gewählt, da eine Eingrenzung auf Titel, Abstract und Schlüsselwörter nur eine geringe Anzahl relevanter Treffer lieferte. In den anderen Datenbanken beschränkte sich die Suche hingegen ausschließlich auf diese drei Elemente. Von den nach Entfernen der Duplikate verbleibenden 746 Publikationen wurden Titel und Abstract gescreent und auf die genannten inhaltlichen Kriterien geprüft. Die danach verbleibenden Artikel wurden anhand der Volltexte untersucht. Da die Ergebnisse der Literatursuche in ein Literaturverwaltungsprogramm importiert worden sind, wurde erst in diesem Schritt deutlich, ob ein Zugriff auf den Volltext über die Lizenzen der HAW Hamburg möglich war. Ein Großteil der nicht zugänglichen Literatur wurde über Scopus identifiziert. Da Scopus lediglich bibliografische Angaben und Abstracts bereitstellt, war ein vorheriges Filtern nach Verfügbarkeit nicht möglich. Nach der Volltextanalyse unter Berücksichtigung der Forschungsfrage blieben 13 relevante Publikationen übrig.

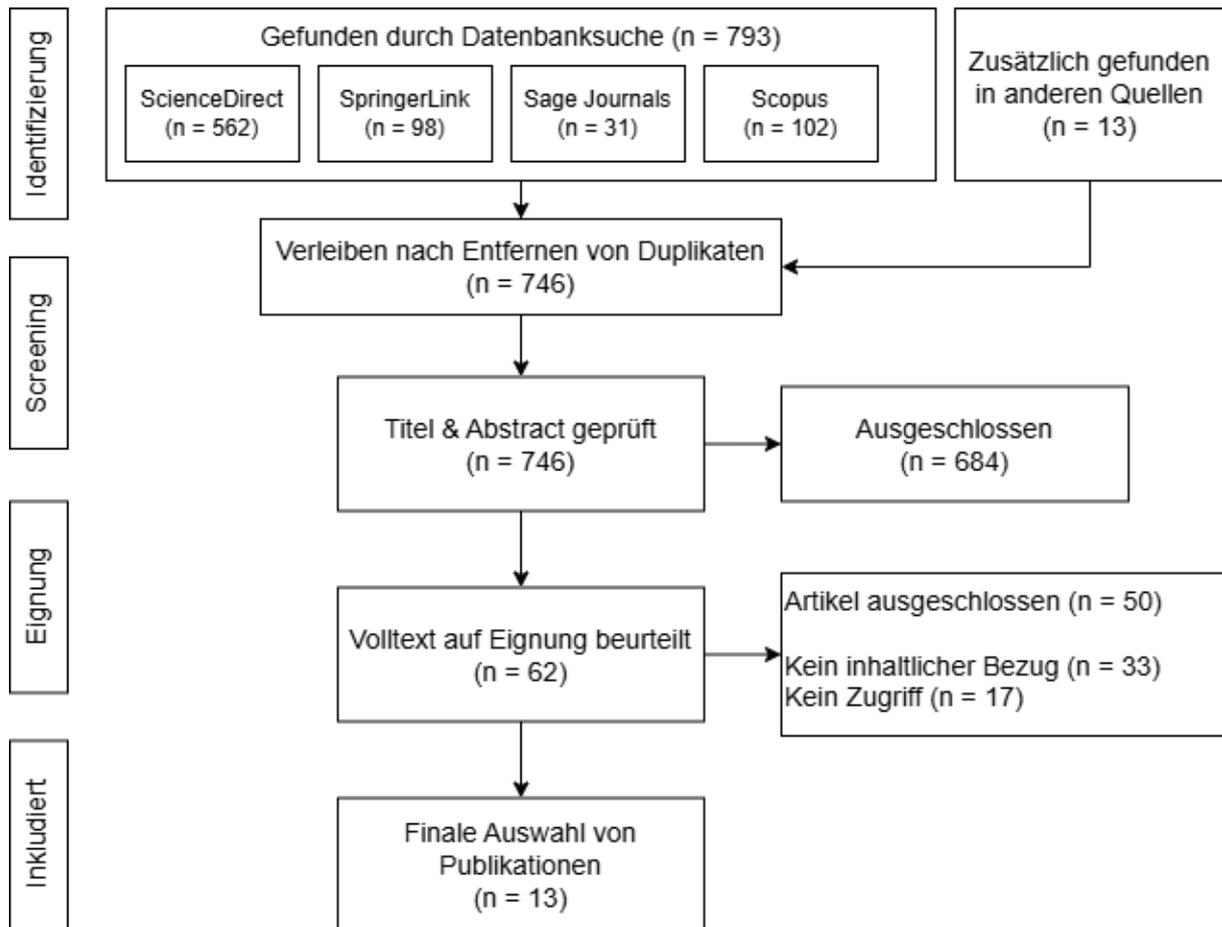


Abbildung 4: Prozess der strukturierten Literaturrecherche

4. Ergebnisse

4.1 Deskriptive Darstellung

Die 13 ausgewählten Publikationen sind in Tabelle 3 aufgeführt. Die grau hinterlegten Veröffentlichungen sind im Rahmen des Petro-HRA Projekts entstanden.

| Jahr | Autor | Titel |
|------|-----------------------|--|
| 2012 | Gould et al. | Human Reliability Analysis in Major Accident Risk Analyses in the Norwegian Petroleum Industry |
| 2014 | Boring | Top-Down and Bottom-Up Definitions of Human Failure Events in Human Reliability Analysis |
| 2015 | Boring | Adapting Human Reliability Analysis from nuclear power to oil and gas applications |
| 2016 | Laumann und Rasmussen | Suggested improvements to the definitions of Standardized Plant Analysis of Risk-Human Reliability Analysis (SPAR-H) performance shaping factors, their levels and multipliers and the nominal tasks |
| 2017 | Rasmussen und Laumann | Decomposition level of quantification in human reliability analysis |
| 2020 | Rasmussen und Laumann | The evaluation of fatigue as a performance shaping factor in the Petro-HRA method |
| 2020 | Taylor et al. | Lessons learned from applying a new HRA method for the petroleum industry |
| 2022 | Bye et al. | The Petro-HRA Guideline, Rev.1, Vol.2 |
| 2014 | Noroozi et al. | Effects of Cold Environments on Human Reliability Assessment in Offshore Oil and Gas Facilities |
| 2018 | Morais et al. | Attempt to predict human error probability in different industry sectors using data from major accidents and Bayesian networks |
| 2020 | Abílio Ramos et al. | A human reliability analysis methodology for oil refineries and petrochemical plants operation: Phoenix-PRO qualitative framework |
| 2020 | Golestani et al. | Human reliability assessment for complex physical operations in harsh operating conditions |
| 2022 | Morais et al. | Robust data-driven human reliability analysis using credal networks |

Tabelle 3: Finale Auswahl der Literatur

4.2 Inhaltliche Darstellung

Die Öl- und Gasindustrie und die Nuklearindustrie eint, dass beide sicherheitskritische Branchen sind, in denen Fehler mit einem hohen Gefährdungspotenzial für Mensch, Technik und Umwelt einhergehen (Boring, 2015, Kapitel 3). Dennoch wurde bereits am Beispiel Norwegens gezeigt, dass die HRA noch kein etabliertes Instrument in der Öl- und Gasindustrie ist und die Betrachtung des menschlichen Einflusses nicht mit dem Niveau in der Nuklearindustrie vergleichbar ist. Auch Boring (2015, Kapitel 2.2) und Ramos (2020, S. 3) führen dies unter anderem auf fehlende regulatorische Anforderungen zurück. Unterschiede in der Regulierung betreffen auch den Ausbildungsstand des Betriebspersonals, das zumindest in amerikanischen

Kernkraftwerken eine staatliche Prüfung bestehen muss (Ramos et al., 2020, S. 3). Darüber hinaus gibt es weitere inhärente Unterschiede zur Nuklearindustrie.

Die Anlagen und Prozesse der Öl- und Gasindustrie sind weitaus vielfältiger und umfassen sowohl die Förderung (Upstream), die Lagerung und den Transport (Midstream) und die Weiterverarbeitung in Raffinerien (Downstream). Auch werden diese Anlagen unter verschiedenen Standortbedingungen, z. B. sowohl On- als auch Offshore, und unter dementsprechend anderen und größeren Umweltrisiken und Einflussfaktoren auf das Betriebspersonal betrieben (Taylor et al., 2020, S. 2). Des Weiteren unterscheiden sich Nutzungsdauer und Entwicklungszyklen von Anlagen, da Kohlenwasserstoffquellen versiegen oder neu erschlossen werden. Daher sind bspw. digitale Mensch-Maschine-Schnittstellen in der Öl- und Gasindustrie weiter verbreitet (Boring, 2015, Kapitel 3).

Die Arbeit in Kernkraftwerken ist deutlich zentralisierter organisiert. Zwar verfügen auch die meisten Anlagen der Öl- und Gasindustrie über einen Kontrollraum, dennoch finden viele Arbeiten außerhalb von diesem statt. Dies geht nach Boring (2015, Kapitel 3) auch mit weniger standardisierten Prozessen und Arbeitsanweisungen einher, die eine flexiblere Reaktion auf äußere Umstände zulassen. Daraus resultiert, dass Handlungen in der Nuklearindustrie laut Boring eher rule-based sind, während diese in der Öl- und Gasindustrie eher skill-based sind. Aus der Diversität der technischen Anlagen und Prozesse resultiert eine größere Anzahl an möglichen Unfallszenarien und unterschiedlichen Verläufen dieser (Ringstad und Szameitat, 2000, nach Gould et al., 2012, S. 2018). Nuklearreaktoren sind typischerweise so entworfen, dass die Anlagen in den ersten 30 min nach einem Vorfall automatisiert laufen und Handmaßnahmen durch das Personal zwar möglich, aber nicht notwendig sind. Dies soll dem Personal Ressourcen zum Verstehen der Situation und Treffen von Entscheidungen geben. Im Gegensatz dazu können sich Vorfälle in Anlagen der Petroleumindustrie deutlich schneller entwickeln und Handlungen durch das Personal innerhalb kürzerer Zeiträume erfordern (Taylor et al., 2020, S. 2; BMUV, 2015, S. 10).

Boring (2015, Kapitel 3) vermutet, dass sich auch die Sicherheitskultur der beiden Industrien unterscheidet bzw. der Variabilität der Sicherheitskultur innerhalb der Öl- und Gasindustrie größer ist. Als mögliche Gründe werden die geografische Abgelegenheit von Offshore-Anlagen und die höhere Fluktuation von Arbeitskräften genannt.

Im Folgenden wird der Umgang mit diesen Unterschieden in den identifizierten Beiträgen beschrieben. Naroozi et al. (2014) und Golestani et al. (2020) betrachten den Einfluss der physischen Umweltbedingungen auf die menschliche Zuverlässigkeit. Anschließend werden mit Petro-HRA, Phoenix-PRO und den Publikationen von Morais et al. Neu- bzw. Weiterentwicklungen von Methoden vorgestellt, die nicht nur einzelne industriespezifische Einflussfaktoren, sondern den gesamten Kontext der Petroleumindustrie betrachten. Abschließend wird auf Basis von Boring kurz die Integration der HRA in die QRA thematisiert.

Einfluss physischer Umweltbedingungen

Noroozi et al. (2014) nutzen HEART zur Quantifizierung des Einflusses von Kälte auf die HEP bei Arbeiten auf Offshore-Anlagen. Niedrige Temperaturen verringern die motorischen Fähigkeiten, Muskelkraft und Mobilität. Schutzkleidung, die diese Auswirkungen vermindern soll, kann hingegen die Mobilität einschränken und die sensorische Wahrnehmung behindern, sodass bspw. auditive Hinweise schlechter wahrgenommen werden. Zur Berücksichtigung dieser Bedingungen schlagen die Autoren vor, in arktischen Bedingungen immer die in Tabelle 3 abgebildeten 95-Perzentile der *Nominal Human Unreliability* der GTT zu nutzen. Zusätzlich wurden 14 der 38 EPCs als besonders relevant für die Arbeit in kalten Umgebungen identifiziert und den Kategorien *physical*, *cognitive*, *instrumentation*, und *management* zugeordnet. In der Berechnung der HEP sollen diese EPCs besonders berücksichtigt werden (Noroozi et al., 2014, S. 826–829). Anhand eines Beispiels wird die HEP unter normalen Bedingungen und kalten Bedingungen berechnet und deren Unterschied mit nichtparametrischen statistischen Tests auf Signifikanz geprüft. Die Autoren kommen zu dem Schluss, dass die Auswirkungen von Kälte signifikant sind (Noroozi et al., 2014, S. 832).

Golestani (2020) berücksichtigt neben dem Einfluss von Kälte auch noch Lärm und Vibration und nutzt ein BN für die Quantifizierung der HEP in zwei Routineszenarien auf Offshore-Anlagen. Die drei Faktoren wurden in direkte Beziehung zu den Human Error Modes gesetzt. Schwellenwerte, bei deren Überschreitung von einem Fehler ausgegangen wird, wurden auf Basis physiologischer Daten festgelegt (Golestani et al., 2020, S. 2 f.). Die Eintrittswahrscheinlichkeiten für Lärm und Vibration wurden statistisch geschätzt, für Kälte basieren die Wahrscheinlichkeiten auf empirischen Daten (Golestani et al., 2020, S. 6). Abbildung 5 zeigt einen Ausschnitt aus diesem BN.

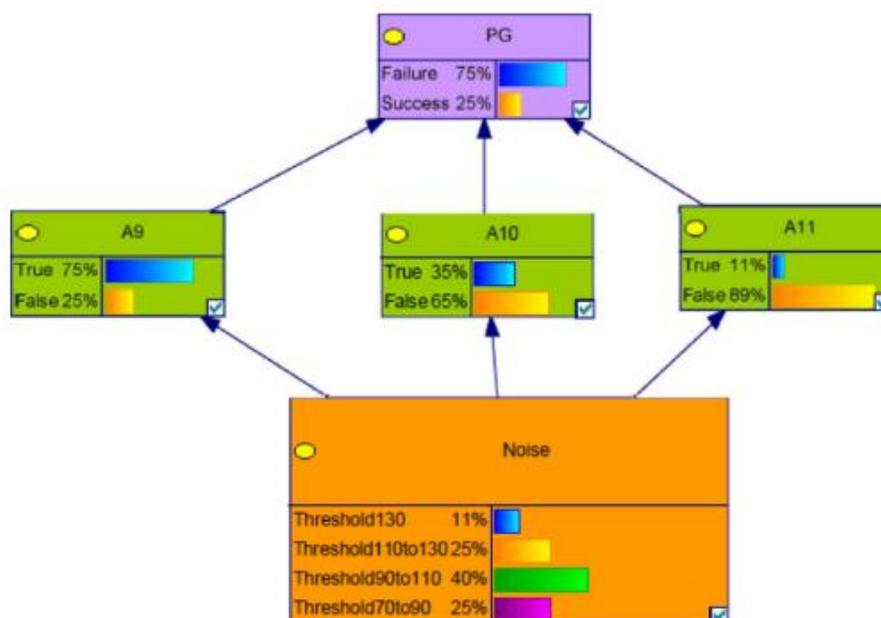


Abbildung 5: Bayes'sches Netzes (Golestani et al., 2020, S. 8)

Neu- und Weiterentwicklungen von HRA-Methoden

Mit Petro-HRA und Phoenix-PRO existieren zwei Methoden, die versuchen, nicht nur einzelne Bedingungen der Öl- und Gasindustrie zu berücksichtigen, sondern laut den Autoren explizit auf die Industrie angepasst worden sind. Im Rahmen des Petro-HRA Projekts wurde SPAR-H als Grundlage gewählt. Als Gründe für diese Methode wurde die Verfügbarkeit von Guidance für die Zuordnung der PSF-Faktoren und Formblättern/Worksheets genannt sowie der im Vergleich zu anderen Methoden geringe Ressourcenaufwand in der Anwendung (Gould et al., 2012, S. 2017). Die vorgenommenen Änderungen in der Quantifizierung betreffen die NHEP und die PSFs. SPAR-H unterscheidet die zwei Aufgabentypen *diagnosis tasks* und *action tasks* und ordnet diesen die NHEP 0,01 bzw. 0,001 zu. In Petro-HRA wird diese Unterscheidung aufgehoben und allen Aufgaben eine NHEP von 0,01 zugeordnet, da die Autoren davon ausgehen, dass jede Aufgabe sowohl kognitive als auch physische Elemente enthält (Laumann & Rasmussen, 2016, S. 298). Die Definitionen/Beschreibungen der PSFs wurden mit dem Ziel angepasst, Überschneidungen zwischen PSFs zu reduzieren und nur PSFs zu berücksichtigen, die einen großen Einfluss haben, tatsächlich bewertbar sind sowie in post-initiator events wirken können. Die Anpassung der Definitionen und quantitativen Faktoren (Multipliers) basierte auf Vergleichen mit anderen HRA-Methoden, Literaturrecherchen und Interviews mit HRA-Anwendern / Experten (Laumann & Rasmussen, 2016, S. 288).

Der PSF *Fitness for Duty / Fatigue* wurde entfernt, da *Fatigue* den Autoren nach ein nicht ausreichend definierter Begriff ist, der Aspekte wie Schlafentzug, Schichtlänge, Nachtschichten und die Dauer der Aufgabenleistung (*Prolonged task performance*) beinhaltet (Rasmussen & Laumann, 2020, S. 3 ff.). Obwohl der Einfluss von *Fatigue* eindeutig nachgewiesen ist, schließen die Autoren den Effekt aus, da dessen Einfluss im Vergleich zu anderen PSFs meist gering ist (Rasmussen & Laumann, 2020, S. 6). Als weiterer Grund wird die durch Experteninterviews identifizierte seltene Nutzung des PSFs genannt (Laumann & Rasmussen, 2016, S. 295). Auch wurde die SPAR-H Definition des PSF *Work Process* als wenig eindeutig und dadurch selten genutzt bewertet. Stattdessen wurden in der Petro-HRA Methode die beiden PSFs *Attitudes to Safety, Work and Management Support* und *Teamwork* neu eingeführt (Laumann & Rasmussen, 2016, S. 296 ff.; Bye et al., 2022, S. 154–162). Des Weiteren wurde der PSF *Ergonomics/HMI* in die beiden PSFs *HMI* und *Physical Working Environment* aufgeteilt. Die Autoren bewerten ergonomische Einflussfaktoren, abgesehen von der Mensch-Maschine-Schnittstelle, innerhalb eines Kontrollraums als wenig relevant für eine HRA. Der PSF *Physical Working Environment* ist daher vor allem für Tätigkeiten außerhalb eines Kontrollraums gedacht (Laumann & Rasmussen, 2016, S. 294 f.). Tabelle 4 stellt die Beschreibungen der Levels und Multiplier des PSF *HMI* von SPAR-H und Petro-HRA gegenüber (Bye et al., 2022, S. 149 f.). Der PSF *HMI* ist der einzige Faktor, dessen Anpassung von den Autoren mit industriespezifischen Bedingungen begründet wurde, da, wie bereits beschrieben, davon ausgegangen wird, dass in der Öl- und Gasindustrie digitale Schnittstellen überwiegen. Alle anderen Änderungen der Definitionen und Multiplier wurden als allgemeingültige Verbesserungsvorschläge präsentiert und daher sind daher an dieser Stelle nicht ausführlicher besprochen (Laumann & Rasmussen, 2016, S. 300)

| SPAR-H | | Petro-HRA | |
|------------|--|------------|---|
| Multiplier | Level Description | Multiplier | Level Description |
| 0,5 | Good: The design of the facility positively impacts task performance, providing needed information and the ability to carry out tasks in such a way that lessens the opportunities for error (e.g., easy to see, use, and understand computer interfaces; instrumentation is readable from workstation location, with measurements provided in the appropriate units of measure). | 0,5 | Low positive effect on performance: The HMI is specifically designed to make human performance as reliable as possible in this task/tasks of this type. |
| 1 | Nominal: The design of the facility supports correct performance, but does not enhance performance or make tasks easier to carry out than typically expected (e.g., operators are provided useful labels; the computer interface is adequate and learnable, although not easy to use). | 1 | Nominal: While the HMI is not specifically designed for making the human performance as reliable as possible for this task/tasks of this type, it corresponds to the stereotypes held by the operators. All of the safety critical information is easily available and no HMI- related issues are interfering with carrying out the task. HMI does not reduce performance nor to a large degree improve performance. |
| 10 | Poor: The design of the facility negatively impacts task performance (e.g., poor labelling, needed instrumentation cannot be seen from a work station where control inputs are made, or poor computer interfaces). | 10 | Moderately negative effect on performance: The HMI causes some problems in either obtaining relevant information or carrying out the task. For example, the HMI does not conform to the stereotypes the operators are used to (e.g., icons, colors, and intuitive placements), or several page changes in the inter-page navigation increases the difficulty in obtaining the required information or carrying out the task. |
| 50 | Missing/Misleading: The required instrumentation fails to support diagnosis or post-diagnosis behavior, or the instrumentation is inaccurate (i.e., misleading). Required information is not available from any source (e.g., instrumentation is so unreliable that operators ignore the instrument, even if it is registering correctly at the time) | 50 | Very high negative effect on performance: The HMI causes major problems in either obtaining relevant information or carrying out the task. For example, the HMI is not designed for the task leading to a difficult work-around, some of the relevant information required for a reliable decision is not made available or, the inter-page navigation creates severe difficulties in obtaining the relevant information or carrying out the task |
| | | P=1 | Extremely high negative effect on performance: A situation where it is not reasonable to assume that the operator/crew will be successful in carrying out the task. An example of this would be a situation where the HMI does not provide the operator/crew with the required information or possibility to perform the task. Alternatively the information provided is misleading to the extent that the operator will not correctly carry out the task. |

Tabelle 4: Gegenüberstellung des PSF HMI aus SPAR-H und Petro-HRA(Bye et al., 2022, S. 149 f.)

Neben der Adaption der Quantifizierung wurde Petro-HRA mit der Absicht entwickelt, eine ausführliche Anleitung/Beschreibung qualitativen Schritte zu bieten, auf Basis derer die Quantifizierung stattfindet. Damit wird einer Empfehlung aus der *International HRA Empirical Study* gefolgt, in welcher fehlende Guidance als Ursache für Variabilität von Ergebnissen identifiziert wird (Forester et al., 2014, S. 89 f.). Petro-HRA wird von den Autoren daher als vollständige (complete) Methode bezeichnet (Taylor et al., 2020, S. 4 f.). Die Auswirkungen unterschiedlichen Vorgehens werden von Rasmussen und Laumann (2017, S. 999 f.) anhand des vereinfachten Beispiels einer Gasleckage gezeigt. In diesem beeinflusst der Grad der Aufgabenzerlegung die HEP. Abbildung 6 zeigt die Ebenen der Aufgabe. Während die Quantifizierung auf Level 1 eher pessimistische HEPs produziert, erschwert die Betrachtung auf einem niedrigeren Level die Zuordnung von PSFs wie Zeit oder Komplexität und ist insgesamt zeitaufwendiger.

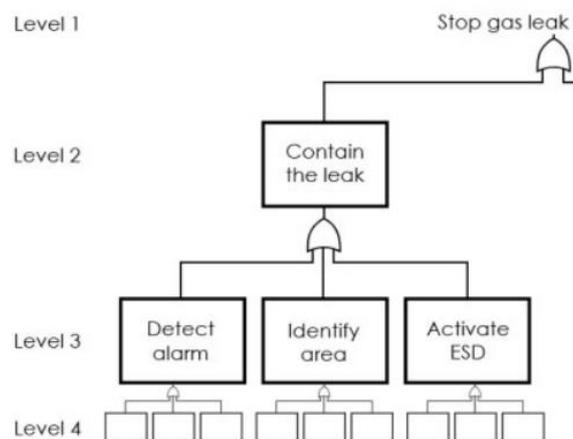


Abbildung 6: Exemplarischer Fehlerbaum einer Gasleckage (Rasmussen und Laumann, 2017, S. 998)

Phoenix for **P**etroleum **R**efining **O**perations basiert auf der für die Nuklearindustrie entwickelten Phoenix-Methode. Phoenix PRO wurde auf Basis von Unfallberichten aus der Petroleumindustrie, Besuchen in Leitwarten von Ölraffinerien, Expertenbefragungen und dem Vergleich zwischen Kontrollräumen und Betrieb von Kernkraftwerken und Raffinerien adaptiert (Ramos et al., 2020, S. 2).

Im Vergleich zu der ursprünglichen Version wurden die in Tabelle 5 dargestellten Crew Failure Modes (CFM), mithilfe derer Phoenix-PRO Fehler klassifiziert, angepasst. Der CFM *Procedure not Used (Intentional)* wurde hinzugefügt, da dieser in analysierten Unfällen ein wichtiger Einflussfaktor war. Der CFM *Decision to Stop Gathering Data* wurde aufgrund der sehr geringen Häufigkeit dieser Fehlerart entfernt (Ramos et al., 2020, S. 6).

In Bezug auf die PIFs wurden *Procedure Updating* und *Knowledge of Plant Conditions* als relevant für den Kontext der Petroleumindustrie erachtet und neu hinzugefügt. Mit dem PIF *Procedure Updating* sollen Situationen berücksichtigt werden, in denen Prozesse nicht der Realität entsprechen, z. B. aufgrund von Änderungen an der Anlage. Der PIF *Time Constraint* wurde in *Stress due to Perceived Situation Urgency* umbenannt. Des Weiteren wurde der PIF *Safety Culture* entfernt, da die Autoren

annehmen, dass dieser u. a. durch die PIFs *Human Machine Interface, Procedure* und *Resources* abgebildet wird (Ramos et al., 2020, S. 8 f.). Im Gegensatz zu Rasmussen und Laumann schätzt Ramos den Einfluss von *Fatigue* als relevant ein und verweist dafür unter anderem auf die Raffinerieexplosion in Texas City in 2005 (Ramos et al., 2020, S. 11). Tabelle 6 stellt alle PIFs von Phoenix Pro dar.

| ID | Crew failure modes in "I" phase | ID | Crew failure modes in "D" phase | ID | Crew failure modes in "A" phase |
|----|--|----|--|----|---|
| I1 | Key Alarm / Information not Responded to (intentional & unintentional) | D1 | Plant/System State Misdiagnosed | A1 | Incorrect Timing of Action |
| I2 | Data Not Obtained (Intentional) | D2 | Procedure Misinterpreted | A2 | Incorrect Operation on Component/Object |
| I3 | Data Discounted | D3 | Failure to Adapt Procedures to the Situation | A3 | Action on Wrong Component/Object |
| I4 | Data Incorrectly Processed | D4 | Procedure Step Omitted (Intentional) | | |
| I5 | Reading Error | D5 | Procedure not Used (Intentional) | | |
| I6 | Information Miscommunicated | D6 | Inappropriate Procedure Followed | | |
| I7 | Wrong Data Source Attended to | D7 | Decision to Delay Action | | |
| I8 | Data Not Checked with Appropriate Frequency | D8 | Inappropriate Strategy Chosen | | |

Tabelle 5: Phoenix-PRO Crew Failure Modes (Ramos et al., 2020, S. 7)

| HMI | Procedures | Resources | Team effectiveness | Knowledge / abilities | Bias | Stress | Task load |
|------------|--|-----------------------------------|---|--|---|---|---|
| HMI Input | Procedure Content | Tools | Communication | Knowledge / Experience / Skill (content) | Motivation/ Commitment | Stress due to Situation Perception | Cognitive Complexity |
| HMI Output | Procedure Updating Procedure Availability | Tool Availability Tool Quality | Communication Quality Communication Availability | Task Training Knowledge of Plant Conditions | Confidence in Instruments Familiarity with or Recency of Situation | Perceived Situation Urgency Perceived Situation Severity | Inherent Cognitive Complexity Cognitive Complexity due to External Factors |
| | | Workplace Adequacy | Team Coordination | Knowledge / Experience / Skill (access) | Competing or Conflicting Goals | Stress due to Decision | Execution Complexity |
| | | | Leadership Team Cohesion | Attention Fitness for Duty | | | Inherent Execution Complexity Execution Complexity due to External Factors |
| | | | Responsibility Awareness Team Composition Team Training | | | | Extra Work Load Passive Information Load |

Tabelle 6: Phoenix-PRO Performance Influencing Factors (Ramos et al., 2020, S. 9)

Morais – Datenbasierte Anwendung eines Credal Networks

Morais et al. (2022, S. 11) nutzen empirische Daten für die Berechnung von HEPs während der Druckentlastung von Lagertanks auf *Floating Production Storage and Offloading Systems* und *Floating Storage Units*. Die qualitative Task-Analyse und die Auswahl der Performance Shaping Factors (PSFs) wurden von Experten durchgeführt, während die Quantifizierung auf Daten des Multi-attributed Technological Accidents-Dataset (MATA-D) basierte (Morais et al., 2022, S. 16). Dieses Datenset umfasst 238 schwerwiegende Unfälle aus verschiedenen Industrien, die nach der von Hollnagel entwickelten CREAM Taxonomie ausgewertet wurden. 46 % der in der Datenbank erfassten Unfälle sind der Öl- und Gasindustrie zugeordnet (Moura et al., 2016, S. 39 f.). Die identifizierten Einflussfaktoren auf ein HFE wurden binär (An- oder Abwesenheit) zugeordnet (Morais et al., 2018, Kapitel 2). Eine Schwäche dieser Daten ist, dass fehlende Kombinationen von PSFs und HFE in BN dazu führen, dass diese Kombinationen als unmöglich betrachtet werden, obwohl sie in Wirklichkeit nur nicht beobachtet wurden. BN erfordern daher eine große Anzahl an Beobachtungen (Morais et al., 2022, S. 4 f.). Diese Lücken werden bisher durch Annahmen von Experten geschlossen, ohne dass die dabei entstehende Unsicherheit sichtbar wird (Morais et al., 2022, S. 5 f.). Credal Networks, eine Erweiterung von Bayessches Netzen, können diese Unsicherheit darstellen, indem die Wahrscheinlichkeit der Variablen durch eine Menge von Wahrscheinlichkeitsverteilungen beschrieben werden, während die Variablen in BN als Punktschätzer dargestellt sind. Der Output wird daher auch in Form eines Wahrscheinlichkeitsintervalls angegeben (Morais et al., 2022, S. 6 f.).

Integration von HRAs in QRAs

Eine HRA kann eigenständig oder im Rahmen eines QRAs durchgeführt werden. Insbesondere in der Nuklearindustrie ist eine HRA in das Probabilistic Safety Assessment (PSA), die industriespezifische Bezeichnung für ein QRA, integriert. Die Verbindung zwischen HRA und PSA ist das HFE. Die in der Nuklearindustrie angewandte Definition für HFEs von der American Society of Mechanical Engineers beschreibt diese als Hardware-Ausfälle, die durch menschliches Handeln oder Nichthandeln ausgelöst werden (ASME, 2009, nach Boring, 2014, S. 563). Diese werden von der PSA vorgegeben. Die HRA identifiziert die genauen Fehlermechanismen und quantifiziert diese. Dieser Ansatz kann als Top-down bezeichnet werden. Im Gegensatz dazu identifiziert der Bottom-up-Ansatz mögliche Fehler und deren Auswirkungen ausgehend von der *task analysis*. Nach Boring (2014, S. 564) ist gerade der Bottom-up-Ansatz in der Öl- und Gasindustrie weitverbreitet. Abbildung 7 verdeutlicht diese unterschiedlichen Vorgehensweisen. Die Herausforderung ist es also, mit der HRA HFEs zu modellieren, die in eine QRA integriert werden können. Boring vermutet, dass der Top-down-Ansatz EOCs weniger zuverlässig aufdeckt als der Bottom-up-Ansatz (Boring, 2014, S. 564).

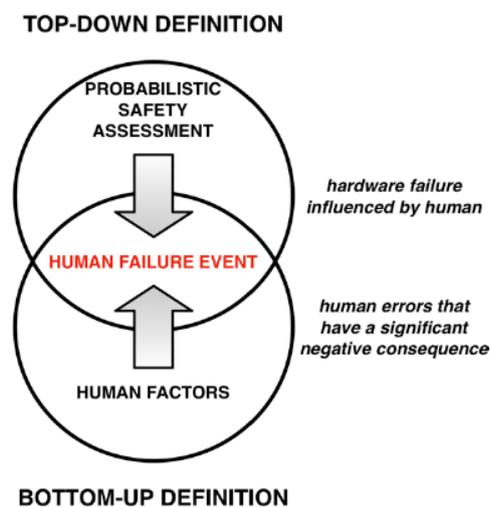


Abbildung 7: Ansätze für die Definition von HFE (Boring, 2014, S. 564)

PETRO-HRA grenzt die möglichen Szenarien und HFEs noch weiter ein, da die Methode nur für die Betrachtung von Post-initiator Szenarien empfohlen ist, d. h. zur Analyse von eindämmenden Handlungen nach Eintritt eines potenziell gefährlichen Ereignisses (Bye et al., 2022, S. 6). Im Rahmen von Phoenix-PRO wird diese Limitierung nicht vorgenommen. Als mögliches mit Phoenix-PRO zu analysierendes Szenario wurde exemplarisch das Wiedereinschalten einer Wärmetauscherbank nach einer Wartung genannt (Ramos et al., 2020, S. 12).

Für die erstmalige Analyse eines Post-initiator Szenarios mit SPAR-H berichten Gould et al. (2012, S. 2017 f.) von einem Zeitaufwand von 100-250 Stunden. Es wird jedoch erwartet, dass sich dieser Aufwand mit der zunehmenden Anzahl durchgeführter Analysen reduziert, da typischerweise Ähnlichkeiten zwischen Systemen und Szenarien bestehen, die genutzt werden können. Die Analyse von Pre-initiator Szenarien wird als deutlich herausfordernder eingeschätzt.

5. Diskussion

5.1 Diskussion der Methode

Die Suchkriterien wurden bewusst sensitiv gewählt, da davon ausgegangen wurde, dass die HRA in der Öl- und Gasindustrie noch nicht etabliert ist. Daher wurde der Suchstring nicht um mögliche weitere Schlüsselwörter, z. B. Namen der verschiedenen HRA-Methoden erweitert, die eine Suche spezifischer gestaltet hätten.

Da die Öl- und Gasindustrie und die Offshore-Industrie als Teil dieser eng mit der maritimen Industrie verzahnt ist, war die Anwendung der Ausschlusskriterien nicht immer eindeutig. So wurde die Publikation von Morais et al. (2022) inkludiert, obwohl ein Szenario auf einem *Floating Storage and Offloading System* betrachtet wurde. Da es sich dabei aber eindeutig um ein Szenario innerhalb der Öl- und Gasindustrie handelte, wurde der Artikel als relevant bewertet. Ausgeschlossen wurden hingegen nautische Szenarien, die bspw. ausschließlich die Bedienung von Dynamischen-Positionier-Systemen betrafen.

Da die Struktur der Quellen sehr unterschiedlich ist, erfolgte die Auswertung nicht einem vorher entwickelten Schema folgend. Die Artikel wurden in Hinblick auf Unterschiede zwischen den Industrien und den daraus resultierenden oder geschlossenen Konsequenzen gelesen und analysiert.

5.2 Diskussion der Ergebnisse

Die Ergebnisse zeigen, dass die industriespezifische Adaptierung von HRA-Methoden auch über die Entwicklung der Petro-HRA Methode hinaus thematisiert wird. Erreicht wird dies insbesondere durch eine Anpassung der PSFs. Mit Phoenix-PRO existiert eine weitere Methode, die den Anspruch vertritt, an die Bedingungen der Öl- und Gasindustrie angepasst zu sein.

Unterschiede zur Nuklearindustrie wurden insbesondere in den mit der Petro-HRA assoziierten Publikationen herausgearbeitet. Während diese sehr schlüssig wirken, wurden sie selten empirisch belegt. Als Beispiel dafür können die Unterschiede in der Variabilität der Sicherheitskultur zwischen den Industrien genannt werden, die von Boring (Boring, 2015, Kapitel 3) als Vermutung formuliert wurden. Auch werden die tatsächlichen Änderungen an der auf SPAR-H basierenden Petro-HRA Methode selten mit den vorher identifizierten Unterschieden, sondern meist mit einer allgemeingültigen, industrieunabhängigen Verbesserung der Nutzerfreundlichkeit begründet (Laumann & Rasmussen, 2016, S. 300). Die in Tabelle 2 dargestellten Änderungen der HMI-Level wurden zwar als industriespezifischen Änderungen dargestellt, können aber auch als allgemeine Anpassung an moderne, digitalisierte Arbeitsumgebungen betrachtet werden und sind vermutlich auch für andere Industrien gültig. Insgesamt sind die veränderten Beschreibungen der PSF-Level wenig spezifisch und nicht mit messbaren Indikatoren verbunden. Einzig der Verweis auf „icons, colors, and intuitive placements“ und „several page changes in the inter-page navigation“ bietet einen spezifischen Anhaltspunkt für die Bewertung digitaler Benutzeroberflächen. Allerdings bietet keine der betrachteten HRA-Methoden messbare Indikatoren zur Beurteilung von PSFs. Die Herausforderung, eine angemessen detaillierte Beschreibung zu formulieren, ohne die Flexibilität eines

Anwenders einzuschränken, wurde auch von den Autoren erkannt und gilt nicht nur für Petro-HRA (Taylor et al., 2020, S. 10). Daraus resultiert aber gleichzeitig die weitere Abhängigkeit von der Erfahrung der Analysten, die in der Öl- und Gasindustrie noch nicht vorhanden ist (Gould et al., 2012, S. 2018). Meiner Meinung nach bietet Petro-HRA als Methode mit umfangreicher Beschreibung der durchzuführenden Schritte daher eher eine Kompensation dieser fehlenden Erfahrung als eine industriespezifische Anpassung. Leider wurden keine Publikationen gefunden, in denen Petro-HRA von Dritten angewendet wird, daher gibt es keine unabhängige Einschätzung der Benutzerfreundlichkeit. Insgesamt dominieren die Publikationen von mit dem Petro-HRA Projekt assoziierten Autoren das Themenfeld.

Petro-HRA und Phoenix-Pro unterschieden sich deutlich im Ablauf. So nutzt Phoenix-PRO ein BN zur Modellierung der Beziehung zwischen PSFs und Fehlern. Daher ist die Anwendung von Phoenix-Pro vermutlich ressourcenintensiver. Darüber hinaus gibt es Unterschiede in der Definition der PSFs. So wird die Wirkung von *Fatigue* unterschiedlich eingeschätzt. Auch wenn die Entscheidung zur In – bzw. Exklusion des PSFs jeweils plausibel begründet ist, wird daraus die Notwendigkeit deutlich, die Modelle und kausalen Beziehungen zwischen Einflussfaktoren und dem menschlichen Verhalten mit wissenschaftlichen Erkenntnissen zu begründen. Hier könnte die HRA von einer engeren Verzahnung mit der HF-Disziplin profitieren.

Die Bandbreite der Möglichkeiten zur Analyse der menschlichen Zuverlässigkeit und die Unterschiede in Komplexität und benötigten Ressourcen zwischen den Methoden wird auch im Vergleich der Publikationen von Noroozi et al. und Golestani et al. deutlich. Die von Noroozi et al. im Abstract gewählte Beschreibung des präsentierten Vorgehens als neue Methode kann infrage gestellt werden. Die Verwendung der 95 Perzentile ist bereits originär in HEART vorgesehen. Die Auswahl der bei Kälte besonders relevanten EPCs kann hilfreich sein, wird aber wenig begründet. Darüber hinaus werden keine Änderungen von HEART vorgeschlagen. Auch verzichten die Autoren darauf, ihr Vorgehen zu diskutieren und Limitationen aufzuzeigen. Im Gegensatz dazu ist die von Golestani et al. präsentierte Anwendung eines BNs deutlich anspruchsvoller.

Differenzen zwischen Titel/Abstract und Volltext erschwerten insgesamt die Identifikation geeigneter Literatur und zeigen die Unterschiede im Niveau der zu dem Thema veröffentlichten Publikationen. So wurden eine Reihe von Publikationen gefunden, die zwar ankündigten, HRA-Methoden im Kontext der Öl- und Gasindustrie zu diskutieren, letztendlich aber nur eine allgemeine Einführung in das Thema darstellten. Exemplarisch kann das Konferenz-Papier *How can the petroleum industry benefit from human reliability analysis?* von Li et al. (2014) genannt werden. In diesem wird einleitend das Ziel des Artikels formuliert: „[...] we provide an overview of the available HRA work and discuss how the HRA methods are applied to or adapted for petroleum industrial operations for managing human errors.“ (Li et al., 2014, S. 2). Tatsächlich wird in dem Paper das generelle Vorgehen einer HRA erklärt und beschrieben. Die angekündigte Diskussion der Adaption beschränkt sich darauf, vereinzelt industriespezifische Beispiele darzustellen. So haben die Autoren die Fehlerklassifikation der ATHAENA-Methode um generische Beispiele aus der Petroleumindustrie erweitert. Auch wenn solche Beispiele sicherlich dabei helfen, das Wissen und Verständnis in Bezug auf HRA-Methoden innerhalb der Öl- und

Gasindustrie zu verbessern, erfüllen diese nicht das einleitend beschriebene Ziel. Ähnliches gilt für das Konferenz-Papier *Tailoring a human reliability analysis to industry needs* von DeMott (2016), welches sich auch darauf beschränkt Vorteile und Grundlagen der HRA zu präsentieren.

Die Vielfalt von Publikationen mit einer großen Anzahl an in die Thematik einführenden Inhalten auf der einen Seite und anspruchsvollen und ressourcenintensiven Methoden wie die Nutzung von Bayesian oder Credal Networks auf der anderen Seite lässt eine Lücke zwischen Praxis und Forschung vermuten. Bei der Auswahl von geeigneten Methoden müssen Anwender daher eine Entscheidung zwischen Aufwand und Aussagekraft wählen.

In der Nuklearindustrie wurde diese Entscheidung durch die zuständigen Aufsichtsbehörden getroffen, die Mindeststandards für eine PSA und HRA festgelegt haben. Doch auch in anderen soziotechnischen Systemen ist es sinnvoll, den menschlichen Einfluss auf die Systemsicherheit in einer Risikoanalyse zu berücksichtigen. Dabei müssen HRA-Methoden vermutlich nicht für jede Industrie neu angepasst werden, solange einerseits Erfahrung und Wissen über HRA-Methoden und andererseits über das betrachtete System vorhanden sind und an einen Tisch gebracht werden.

Im Rahmen der Energiewende werden weiterhin Systeme entstehen, bei denen die Einbeziehung des menschlichen Faktors in Risikoanalysen sinnvoll ist. Auch wenn neue Anlagen, wie beispielsweise Windenergie- oder Solaranlagen im laufenden Betrieb weitgehend autonom arbeiten, entstehen durch diese Anlagen neue Herausforderungen, etwa die Steuerung eines volatilen Stromnetzes. Selbst wenn dieses größtenteils automatisch geregelt würde, wird es absehbar weiterhin Menschen in Schaltwarten geben, die mindestens eine überwachende Funktion übernehmen. In unseren modernen Systemen sind daher die anfänglich erwähnten *Ironien der Automatisierung* auch mehr als 40 Jahre nach der Beschreibung durch Bainbridge noch äußerst relevant und eines von vielen Argumenten für die Notwendigkeit eines ganzheitlichen Risikomanagements. Auch lassen sich die beschriebenen Erkenntnisse über den Einfluss und die Bewertung von widrigen Umweltbedingungen auf bspw. (Wartungs-)Arbeiten in Offshore-Windparks übertragen.

So wie sich das von soziotechnischen Systemen ausgehende Risiko nicht vollständig vermeiden lässt, wird sich auch die mit einer Risikoanalyse verbundene Unsicherheit der Ergebnisse nie vollständig eliminieren lassen. Solange die Anwender dieser Schwächen bewusst sind und keine Alternativen vorhanden sind, bieten HRA-Methoden eine bewährte und angemessene Möglichkeit, Risiken einzuschätzen. Erik Hollnagel hat dazu treffend formuliert:

[...] As I see it, a method is always developed to deal with the problems at the time. (This also goes for the more traditional analysis methods that are linked to a specific model.) Since the problems will change, any method – [...] – will sooner or later become outdated and obsolete. No one can say when that will happen, but when it happens there will surely be something to take its place. [...]. (Patriarca, Di Gravio, et al., 2020, S. 19).

Literaturverzeichnis

Allgemeine Bundesbergverordnung vom 23. Oktober 1995 (BGBl. I S. 1466), die zuletzt durch Artikel 4 der Verordnung vom 18. Oktober 2017 (BGBl. I S. 3584) geändert worden ist

Badke-Schaub, P., Hofinger, G., & Lauche, K. (Hrsg.). (2012). Human Factors: Psychologie sicheren Handelns in Risikobranchen ; mit 17 Tabellen (2., überarb. Aufl). Springer.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6), 775–779.
[https://doi.org/10.1016/0005-1098\(83\)90046-8](https://doi.org/10.1016/0005-1098(83)90046-8)

BMUV. (2015). Sicherheitsanforderungen an Kernkraftwerke (Version 02/22). Verfügbar unter: <https://www.bmuv.de/download/sicherheitsanforderungen-an-kernkraftwerke>

Boring, R. L. (2007). Dynamic human reliability analysis: Benefits and challenges of simulating human performance. In European Safety and Reliability Association (Hrsg.), *Proceedings of the European Safety and Reliability Conference 2007, ESREL 2007—Risk, Reliability and Societal Safety* (S. 1043–1049). Taylor & Francis/Balkema. Verfügbar unter: https://www.researchgate.net/publication/281963776_Dynamic_human_reliability_analysis_Benefits_and_challenges_of_simulating_human_performance

Boring, R. L. (2012). Fifty years of THERP and human reliability analysis. 11th international probabilistic safety assessment and management conference and the annual European safety and reliability conference 2012, Volume 5, 3523–3532. Verfügbar unter: <https://inldigitallibrary.inl.gov/sites/sti/sti/5680968.pdf>

Boring, R. L. (2014). Top-Down and Bottom-Up Definitions of Human Failure Events in Human Reliability Analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 563–567.
<https://doi.org/10.1177/1541931214581119>

- Boring, R. L. (2015). Adapting Human Reliability Analysis from nuclear power to oil and gas applications. Proceedings of the 25th European Safety and Reliability Conference, ESREL 2015, 2853–2860. Verfügbar unter: https://www.researchgate.net/publication/281704363_Adapting_Human_Reliability_Analysis_from_nuclear_power_to_oil_and_gas_applications
- Boring, R. L., Roth, E., Straeter, O., Laumann, K., Blackman, H. S., Oxstrand, J., & Persensky, J. J. (2009). Is Human Reliability Relevant to Human Factors?. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 53(10), 610–614. <https://doi.org/10.1177/154193120905301006>
- Bundesberggesetz vom 13. August 1980 (BGBl. I S. 1310), das zuletzt durch Artikel 4 des Gesetzes vom 22. März 2023 (BGBl. 2023 I Nr. 88) geändert worden ist
- Bundes-Immissionsschutzgesetz in der Fassung der Bekanntmachung vom 17. Mai 2013 (BGBl. I S. 1274; 2021 I S. 123), das zuletzt durch Artikel 1 des Gesetzes vom 3. Juli 2024 (BGBl. 2024 I Nr. 225) geändert worden ist
- Bye, A., Laumann, K., Blackett, C., Rasmussen, M., Øie, S., Merwe, K. van de, Øien, K., Boring, R., Paltrinieri, N., Wærø, I., Massaiu, S., Gould, K., Farbrot, J. E., & Fernander, M. (2022). The Petro-HRA Guideline, Rev.1, Vol.2 (IFE/E-2022/002). Institute for Energy Technology. Verfügbar unter: <https://ife.brage.unit.no/ife-xmlui/handle/11250/2989744>
- National Commission on the BP Deepwater Horizon Oil Spill and Offshore Drilling. (2011). Deep Water: The Gulf Oil Disaster And The Future Of Offshore Drilling - Report to the President (BP Oil Spill Commission Report). Verfügbar unter: <https://www.govinfo.gov/app/details/GPO-OILCOMMISSION>
- DeMott, D. L. (2016). Tailoring a human reliability analysis to industry needs. 2016 Annual Reliability and Maintainability Symposium (RAMS), Tucson, AZ, USA, pp. 1-4. <https://doi.org/10.1109/RAMS.2016.7448030>
- Forester, J., Dang, V. N., Bye, A., Lois, E., Massaiu, S., Broberg, H., Braarud, P. Ø., Boring, R., Männistö, I., Liao, H., Julius, J., Parry, G., & Nelson, P. (2014). The International HRA Empirical Study—Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data (NUREG-2127). U.S.

- Nuclear Regulatory Commission. Verfügbar unter: <https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr2127/index.html>
- Golestani, N., Abbassi, R., Garaniya, V., Asadnia, M., & Khan, F. (2020). Human reliability assessment for complex physical operations in harsh operating conditions. *Process Safety and Environmental Protection*, 140, 1–13. <https://doi.org/10.1016/j.psep.2020.04.026>
- Gould, K. S., Ringstad, A. J., & Van De Merwe, K. (2012). Human Reliability Analysis in Major Accident Risk Analyses in the Norwegian Petroleum Industry. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 2016–2020. <https://doi.org/10.1177/1071181312561421>
- Hou, L.-X., Liu, R., Liu, H.-C., & Jiang, S. (2021). Two decades on human reliability analysis: A bibliometric analysis and literature review. *Annals of Nuclear Energy*, 151, 107969. <https://doi.org/10.1016/j.anucene.2020.107969>
- International Atomic Energy Agency. (2016). *Safety Assessment for Facilities and Activities*. Verfügbar unter: <https://www.iaea.org/publications/10884/safety-assessment-for-facilities-and-activities>
- International Atomic Energy Agency. (2024). *Development and Application of Level 1 Probabilistic Safety Assessment for Nuclear Power Plants (Rev. 1)*. <https://doi.org/10.61092/iaea.3ezv-lp49>
- Kommission für Anlagensicherheit. (2008). *Leitfaden: Empfehlungen für interne Berichtssysteme als Teil des Sicherheitsmanagementsystems gemäß Anhang III Störfall-Verordnung (KAS-8)*. Verfügbar unter: <https://www.kas-bmu.de/kas-leitfaeden-arbeits-und-vollzugshilfen.html>
- Kommission für Anlagensicherheit. (2011). *Leitfaden Kompetenzen bezüglich menschlicher Faktoren im Rahmen der Anlagensicherheit (Betreiber, Behörden und Sachverständige) (KAS-20)*. Verfügbar unter: <https://www.kas-bmu.de/kas-leitfaeden-arbeits-und-vollzugshilfen.html>
- Kommission für Anlagensicherheit. (2018). *Leitfaden zum Konzept zur Verhinderung von Störfällen und zum Sicherheitsmanagementsystem (KAS-19)*. Verfügbar unter: <https://www.kas-bmu.de/kas-leitfaeden-arbeits-und-vollzugshilfen.html>

- Kirwan, B. (2015). Human Reliability Assessment. In J. R. Wilson & S. Sharples (Hrsg.), *Evaluation of human work* (Fourth edition, S. 793–822). CRC Press, Taylor & Francis Group.
- Laumann, K., Blackman, H., & Rasmussen, M. (2018). Challenges with data for human reliability analysis. In Haugen et al. (Hrsg.), *Safety and Reliability – Safe Societies in a Changing World*, 315–322. Taylor & Francis Group. Verfügbar unter: <https://www.taylorfrancis.com/chapters/oa-edit/10.1201/9781351174664-40/challenges-data-human-reliability-analysis-laumann-blackman-rasmussen>
- Laumann, K., & Rasmussen, M. (2016). Suggested improvements to the definitions of Standardized Plant Analysis of Risk-Human Reliability Analysis (SPAR-H) performance shaping factors, their levels and multipliers and the nominal tasks. *Reliability Engineering and System Safety*, 145, 287–300.
<https://doi.org/10.1016/j.ress.2015.07.022>
- Li, Y., Cho, J. J., & Ren, Y. (2014). How can the petroleum industry benefit from human reliability analysis?. *IADC/SPE Drilling Conference and Exhibition*, 1, 569–582. Verfügbar unter:
https://www.researchgate.net/publication/267459480_How_Can_the_Petroleum_Industry_Benefit_From_Human_Reliability_Analysis
- Liu, P., Zhang, R., Yin, Z., & Li, Z. (2021). HUMAN ERRORS AND HUMAN RELIABILITY. In G. Salvendy & W. Karwowski (Hrsg.), *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS* (1. Aufl., S. 514–572). Wiley.
<https://doi.org/10.1002/9781119636113.ch20>
- Mkrtchyan, L., Podofillini, L., & Dang, V. N. (2015). Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability Engineering & System Safety*, 139, 1–16.
<https://doi.org/10.1016/j.ress.2015.02.006>
- Morais, C., Estrada-Lugo, H. D., Tolo, S., Jacques, T., Moura, R., Beer, M., & Patelli, E. (2022). Robust data-driven human reliability analysis using credal networks. *Reliability Engineering and System Safety*, 218.
<https://doi.org/10.1016/j.ress.2021.107990>

- Morais, C., Moura, R., Beer, M., & Patelli, E. (2018). Attempt to predict human error probability in different industry sectors using data from major accidents and Bayesian networks. In 14th Probabilistic Safety Assessment and Management: PSAM 2018, 2018-09-16 - 2018-09-21. Verfügbar unter: https://www.researchgate.net/publication/328513274_Attempt_to_predict_human_error_probability_in_different_industry_sectors_using_data_from_major_accidents_and_Bayesian_networks
- Moura, R., Beer, M., Patelli, E., Lewis, J., & Knoll, F. (2016). Learning from major accidents to improve system design. *Safety Science*, 84, 37–45. <https://doi.org/10.1016/j.ssci.2015.11.022>
- Noroozi, A., Abbassi, R., MacKinnon, S., Khan, F., & Khakzad, N. (2014). Effects of Cold Environments on Human Reliability Assessment in Offshore Oil and Gas Facilities. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(5), 825–839. <https://doi.org/10.1177/0018720813512328>
- NORSOK Standard Z-013 Risk and emergency preparedness assessment. Edition 3, October 2010. Standards Norway
- Offshore-Bergverordnung vom 3. August 2016 (BGBl. I S. 1866), die durch Artikel 12 der Verordnung vom 29. November 2018 (BGBl. I S. 2034; 2021 I 5261) geändert worden ist
- Patriarca, R., Di Gravio, G., Woltjer, R., Costantino, F., Praetorius, G., Ferreira, P., & Hollnagel, E. (2020). Framing the FRAM: A literature review on the functional resonance analysis method. *Safety Science*, 129, 104827. <https://doi.org/10.1016/j.ssci.2020.104827>
- Patriarca, R., Ramos, M., Paltrinieri, N., Massaiu, S., Costantino, F., Di Gravio, G., & Boring, R. L. (2020). Human reliability analysis: Exploring the intellectual structure of a research field. *Reliability Engineering & System Safety*, 203, 107102. <https://doi.org/10.1016/j.ress.2020.107102>
- Ramos, M. A., López Droguett, E., Mosleh, A., & Das Chagas Moura, M. (2020). A human reliability analysis methodology for oil refineries and petrochemical plants operation: Phoenix-PRO qualitative framework. *Reliability Engineering and System Safety*, 193. <https://doi.org/10.1016/j.ress.2019.106672>

- Ramos, M. A., Patriarca, R., & Paltrinieri, N. (2022). The social structure of 70 years of literature on Human Reliability. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 236(6), 911–922.
<https://doi.org/10.1177/1748006X211053122>
- Rasmussen, M., & Laumann, K. (2017). Decomposition level of quantification in human reliability analysis. In Walls, Revie & Bedford (Hrsg.), *Risk, Reliability and Safety: Innovating Theory and Practice*, 997–1002. Taylor & Francis Group. Verfügbar unter:
https://www.researchgate.net/publication/308969723_Decomposition_level_of_quantification_in_human_reliability_analysis
- Rasmussen, M., & Laumann, K. (2020). The evaluation of fatigue as a performance shaping factor in the Petro-HRA method. *Reliability Engineering and System Safety*, 194. <https://doi.org/10.1016/j.ress.2018.06.015>
- Richtlinie 2012/18/EU des Rates vom 4. Juli 2012 zur Beherrschung der Gefahren schwerer Unfälle mit gefährlichen Stoffen. Verfügbar unter: <https://eur-lex.europa.eu/eli/dir/2012/18/oj>
- Rochette, J. (2014). International regulation of offshore oil and gas activities: Time to head over the parapet (Policy Brief 06/14). IDDRI. Verfügbar unter:
https://www.iddri.org/sites/default/files/import/publications/pb0614_jr_offshore_en.pdf
- Selvik, J. T., & Bellamy, L. J. (2020). Addressing human error when collecting failure cause information in the oil and gas industry: A review of ISO 14224:2016. *Reliability Engineering & System Safety*, 194, 106418.
<https://doi.org/10.1016/j.ress.2019.03.025>
- Skogdalen, J. E., & Vinnem, J. E. (2011). Quantitative risk analysis offshore - Human and organizational factors. *Reliability Engineering & System Safety*, 96(4), 468–479. <https://doi.org/10.1016/j.ress.2010.12.013>
- Störfall-Verordnung in der Fassung der Bekanntmachung vom 15. März 2017 (BGBl. I S. 483), die zuletzt durch Artikel 7 des Gesetzes vom 3. Juli 2024 (BGBl. 2024 I Nr. 225) geändert worden ist

Taylor, C., Øie, S., & Gould, K. (2020). Lessons learned from applying a new HRA method for the petroleum industry. *SI:HRA FOUNDATIONS & FUTURE*, 194, 106276. <https://doi.org/10.1016/j.res.2018.10.001>