

7 Grundlagen des diagnostischen Prozesses und der diagnostischen Urteilsbildung

Die Merkmale des diagnostischen Prozesses sind:

- Die zeitliche, organisatorische, strategische und personale Erstreckung zwischen vorgegebenen, zunächst allgemeinen und später präzisierten diagnostischen Fragestellungen sowie deren Beantwortung.
- Die Beantwortung erfolgt in Form einer Diagnose oder Prognose.
- Im Verlauf des diagnostischen Prozesses können aus der Beantwortung einer Teilfragestellung neue Fragen entstehen, so dass eine Verschachtelung von Fragen und Beantwortungen daraus hervorgeht.
- Die Beantwortung der Fragestellung ist zugleich Ausgangspunkt für eine Entscheidungshilfe und Entscheidungen für den Auftraggeber.

Der diagnostische Prozess ist wie eine empirische Untersuchung.

Gemeinsamkeiten

1. Fragestellung und Übersetzung in psychologische Terminologie
2. Fragestellung und Vorinformationen bestimmen die Aufstellung einer Hypothese
3. Überprüfung der Hypothese mit geeigneten diagnostischen Mitteln

Unterschiede

- Die Daten sollen grundsätzlich nach dem Prinzipien der Sparsamkeit und Eignung erhoben werden
- Empirische Studien werden oft an großen Stichproben erhoben, der diagnostische Prozess betrifft aber häufiger nur eine Person
- Es geht im diagnostischen Prozess also um die Überprüfung idiographischer Hypothesen
- Ein diagnostischer Prozess kann mehrere Teiluntersuchungen umfassen, anders als eine empirische Studie, in der die Fragestellung a priori

festgelegt wird und weitere daraus resultierende Fragen in einer neuen Studie erforscht werden müssen.

- Bei der Beantwortung der ursprünglichen Ausgangsfrage, muss die Antwort wieder in die Sprache des Auftraggebers zurückübersetzt werden und mündlich oder schriftlich kommuniziert werden.

In der Pflichtliteratur wird der diagnostische Prozess in Rahmen eines handlungstheoretischen Modells von Kaminski (1970) als Arbeitsprozess mit einer Rückmeldungsschleife dargestellt, der erst mit Erreichen des Zielkriteriums abgeschlossen ist. Der Arbeitsprozess umfasst folgende Phasen

1. Phase: Eingangsdaten
2. Phase: Hypothesenbildung und –bearbeitung
3. Phase: Hypothesenbeurteilung
4. Phase: Planung der Datenbeschaffung
5. Phase: Planung des praktischen Teils
6. Phase: Praktische Umsetzung
7. Phase: Prüfung der Praxis

Ein besonders zentraler und kritischer Aspekt des diagnostischen Prozesses ist die diagnostische [Urteilsbildung](#)

In der zunächst folgenden Pflichtliteratur wird die Urteilsbildung unterschieden nach:

- [Klinischer Urteilsbildung](#)
- [Statistischer Urteilsbildung](#)

Pflichtliteratur, Krohne und Hock (2007), Psychologische Diagnostik und Anwendungsfelder. Kapitel 6: Der Prozess der diagnostischen Urteilsbildung

Diagnostische Urteile sind Aussagen, die auf der Basis vorliegender oder eigens erhobener, psychologischer Daten über eine Person, eine Gruppe oder einen Sachverhalte getroffen werden.

Die Urteilsbildung ist Teil eines mehrstufigen und rückgekoppelten Vorganges (siehe Skript und Pflichtliteratur)

Zentrale Komponente der Urteilsbildung
Datenkombination oder –integration

Drei Fragestellungen

1. Wie kommt ein diagnostisches Urteil zustande?
2. Wie gültig sind diese Urteile?
3. Wie lässt sich deren Qualität sichern ggf. optimieren?

Ausgangspunkt der Autoren ist die zweite Frage; hier wird die Validität der Urteilsbildung angesprochen.

Es gab lang anhaltende Diskussionen über die Validität und deren Vor- und Nachteile **klinischer (informeller)** und **statistischer (formeller)** Methoden der Datenkombination.

Gewichtung und Kombination von Daten lassen sich mit Hilfe so genannter **paramorpher** Modelle abbilden. Es handelt sich um Modelle, in denen die Datenkombination bei Diagnostikern in formeller Weise dargestellt wird. Mit diesen Modellen wird die Urteilsbildung explizit beschrieben. Paramorphe Modelle liefern auch Anhaltspunkte über die Verbesserung des diagnostischen Vorgehens.

6.1 Klinische und statistische Urteilsbildung

Historisch ging es zunächst nur um die Validität also:

- Wie gut treffen Diagnose und Vorhersage zu
- Die mittels unterschiedlicher Methoden gewonnen wurden

In den 1920er Jahren gibt es zwei Positionen zum Verfahren der Diagnosefindung

1. Klinisches Verfahren
2. Statistisches Verfahren

Prognostische Hauptfragestellungen (die Debatte wurde unter dem Titel „klinische vs. statistische Vorhersagen“ geführt, aber generell es geht nicht nur um Prognosen)

- Erfolgchancen bei bestimmten Therapien
- Rückfallrisiko von Straftätern
- Akademischer Erfolg von Studienbewerbern
- Ausgang schwerwiegender physischer Eingriffe bei psychischen Störungen

Klinische Vorhersage

- Standardmodell
- Grundlage ist das menschliche Beurteilen
- Beurteilung erfolgt durch Experten oder –gremien, Psychokongressen
- Orientierung an einem konkret vorliegenden Fall
- Manifestierung in der Bevorzugung biographischer Formen der Datenerhebung (Gespräch, Anamnese, Interview usw.) und Datenintegration
- Die Datenintegration ist kasuistisch → Diagnosen werden durch Vergleich zu ähnlichen Fällen, bzw. den spezifischen Differenzen zum aktuellen Fall gestellt
- Klinische Vorhersage ist erfahrungs- und intuitionsgesteuert
- Die Regeln, aufgrund derer Entscheidungen getroffen werden, sind nicht explizit und schon gar nicht empirisch validiert (sagen Kritiker)

- Die lange Erfahrung, mehrjährige Ausbildung und Fähigkeit große Datenmenge zu integrieren, sollte aber den Diagnostiker zu validen Aussagen befähigen.
- Wird nicht nur in der klinisch psychologischen Diagnostik verwendet sondern auch in anderen Bereichen der Angewandten Psychologie z.B. der Personalauslese.

Statistische Vorhersage

- Stützen sich allein auf empirisch gesicherte Regelmäßigkeit der Daten für einen vorliegenden Fall
- Es liegen bereits Zusammenhänge zwischen Prädiktor- und Kriteriumsvariablen vor, die auf einen neuen Fall angewendet werden können und damit in die Zukunft projizierbar sind
- Prädiktorvariablen sind Variablen mit bekannten Ausprägungen. Sie werden zur Vorhersage noch unbekannter Kriterien herangezogen
 - Beispiel, Prädiktoren bei Straftätern:
Schwere der Tat
Häufigkeit der Rückfälle bei solchen Taten
Das Geschlecht
 - Beispiel, Akademischer Erfolg:
Allgemeine Intelligenz
Schulabschlussnote
- Ein interessierendes Kriterium wird durch eine gewichtete Kombination von Prädiktoren vorhergesagt
- Regeln sind explizit, keine menschliche Beurteilung
- Menschliche Beurteilung kann bei der Datenerhebung eine Rolle spielen, z.B. bei der Beurteilung der Schwere einer Straftat
- Individuelle Besonderheiten werden nur berücksichtigt, wenn sie durch Prädiktoren, die in das Modell miteinbezogen abgedeckt sind
- Die Menge der Prädiktoren ist begrenzt
- Es fallen viele Eigenheiten konkreter Fälle unter den Tisch

Empirische Untersuchungen darüber welche der beiden Vorgehensweisen bessere Ergebnisse liefert, ließen bei aller hitzigen Debatten keine akzeptierten Schlussfolgerungen zu.

Meehl der 1954 ein Buch über diese Problematik veröffentlichte, wies darauf hin, dass es sich bei der Datenerhebung (oder –messung) und der Datenkombination um zwei logisch von einander zu unterscheidende Schritte handele. Wer die Leistungsfähigkeit einer diagnostischen Prozedur untersuchen wolle, müsse beide Vorgänge separat betrachten.

Datenerhebung → psychometrisch/nichtpsychometrisch

↓

Datenkombination → Formell-statistisch/informell-klinisch

↓

Diagnose

Vorhersage

Entscheidung

Psychometrische Daten

Daten werden von Test und Verfahren geliefert, die standardisiert sind und eine einheitliche Klassifikation und Verrechnung der Reaktionen gewährleisten. Diese Daten unterliegen nicht dem Ermessen einer urteilenden Person.

Nichtpsychometrische Daten

Eindrücke eines Diagnostikers, Äußerungen, bzw. Erleben und Verhalten einer Person betreffend.

Der beurteilende Diagnostiker ist hier das Messinstrument.

Die Unterscheidung psychometrisch, nicht psychometrisch ist nicht gleich quantitativ und qualitativ. Man kann auch nicht psychometrische Daten quantitativ erheben. Auch Objektivität und Zuverlässigkeit sind nicht jeweils mehr oder weniger gegeben.

Kombinationen psychometrischer und nichtpsychometrischer mit formellen und informellen Daten

1. **Nichtpsychometrische** Daten werden **informell** kombiniert
Beispiel: Fahrtüchtigkeit eines alkoholauffälligen Verkehrsteilnehmers, beurteilt durch ein Interview
2. **Psychometrische** Daten werden **informell** kombiniert
Beispiel: Zur Diagnose werden Ergebnisse aus Konzentrationstests genutzt
3. **Nichtpsychometrische** Daten werden **formell** kombiniert
Beispiel: Durchführung einer Interviews, die Daten daraus werden regelgeleitet integriert
4. **Psychometrische** Daten werden **formell** kombiniert
Beispiel: Ergebnisse eines Konzentrationstests werden in eine Gleichung zur Bestimmung des Unfallrisikos eingesetzt.
5. **Psychometrische und nichtpsychometrische** Daten werden **informell** oder **formell** kombiniert. Das sind die häufigsten Fälle in der diagnostischen Praxis.

Die von Meehl beklagte Konfundierung entsteht deshalb, weil die Befürworter formeller Methoden häufig psychometrische (harte) Verhaltensdaten bevorzugen und die Befürworter informeller Methoden nichtpsychometrischen (weichen) Daten den Vorzug geben.

6.1.2 Statistische Vorhersage

Für Vorhersagezwecke ist die statistische Methode zunächst die naheliegende.

Häufiges Modell → multiple Regression

Die Variation einer quantitativen Kriteriumsvariablen wird durch additive und gewichtete Kombination mehrerer quantitativer oder binärer Prädiktorvariablen beschrieben.

Beispiel: Angenommen das Kriterium sei das Maß für den Berufserfolg (Gehaltsstufe, Durchschnittsnote, akademische Ausbildung, oder ein Indikator

für Lebenszufriedenheit nach einer therapeutischen Behandlung). Prädiktoren wären dann Daten aus Fähigkeitstests, Interessentests oder Persönlichkeitsinventaren.

Voraussetzung für eine regressionsanalytische Vorhersage ist, dass sich das Kriterium in *einem* quantitativen Wert ausdrücken lässt, dafür können auch mehrere Variablen aggregiert (zusammengefasst) werden.

Liegen Messungen und Werte für beide Variablen (Kriterium Y, Prädiktoren X_1, X_2, X_3, \dots) vor, lässt sich das Regressionsmodell durch eine Gleichung ausdrücken:

$$\hat{Y} = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Y-Dach repräsentiert den Vorhersagewert des Kriteriums, a ist die additive Konstant, b_1 bis b_n sind die Gewichte, mit denen die Prädiktoren zur Vorhersage beitragen. Sie heißen Regressionsgewichte. In der Regressionsanalyse werden die Parameter (a und b) so festgelegt, dass mathematisch eine optimale Vorhersage möglich ist, nämlich die quadrierten Abweichungen zwischen Y-Dach und den faktischen Kriteriumswerten Y minimal werden. Wie genau sich nun Y durch \hat{Y} vorhersagen lässt, lässt sich durch eine Korrelation zwischen beiden ausdrücken. Dieses Maß der multiplen Korrelation ist besonders interessant und wird durch (groß) R ausgedrückt. Es kennzeichnet den Zusammenhang zwischen mehreren Prädiktoren und einem Kriterium und kann prinzipiell zwischen 0 und 1 liegen.

Wert nahe Null = Schwacher Zusammenhang

Werte nahe Eins = Starker Zusammenhang

R^2 , das Quadrat der multiplen Korrelation, gibt den Anteil der Varianz des Kriteriums an, das durch alle im Modell enthaltenen Prädiktoren aufgeklärt wird.

Wird ein bedeutsamer Zusammenhang gefunden, ist natürlich die Frage, welcher der Prädiktoren mit welchem Gewicht zur Vorhersage beiträgt. Dafür werden die Regressionskoeffizienten (b – Gewichte) herangezogen. Diese Gewichte geben an, um welchen Betrag sich das Kriterium erhöht,

wenn ein Prädiktor um eine Skaleneinheit erhöht wird, während alle anderen Prädiktoren konstant bleiben.

Zum besseren Vergleich, z.B. bei IQ-Werten, werden die Regressionsgewichte standardisiert. Man kann sie z-transformieren, die daraus entstehenden Gewichte, werden als β – Gewichte bezeichnet. Die Interpretation ist analog zu den b-Gewichten nur dass es sich hier um Einheiten der Standardabweichung von Kriterium und Prädiktoren handelt.

Bei der einfachen (nicht multiplen) Regression sind die standardisierten Gewichte identisch mit der Korrelation zwischen Prädiktor und Kriterium.

Generell soll viel Varianz im Kriterium mit so wenig Prädiktoren wie möglich aufgeklärt werden. Komplexität wird dann akzeptiert, wenn sich dadurch die Aufklärung des Kriteriums durch Hinzunehmen weiterer Prädiktoren verbessern lässt.

Ist ein passendes statistisches Modell gefunden, sollte eine Kreuzvalidierung vorgenommen werden, die das Modell anhand neuer Daten überprüft. Die, anhand einer Stichprobe A, gewonnenen Gewichte, werden zur Vorhersage des Kriteriums bei einer zweiten Stichprobe B verwendet, deren Daten wiederum nicht bereits in die Berechnung eingegangen sein dürfen. Die Anwendung des Modells auf Daten einer anderen Stichprobe führt zu einem geringeren, aber realistischeren Schätzwert der Vorhersagekraft. Es wird zusätzlich geprüft ob sich die Gewichte, die den einzelnen Prädiktorvariablen zugeordnet wurden, auch in einer neuen Stichprobe bewähren.

Die Regression ist ein additives Modell, d.h. geringe Werte bei einem Konstrukt können durch hohe Werte eines anderen kompensiert werden, z.B. geringes räumliches Vorstellungsvermögen und hohe mathematische Kenntnisse bei Tests zu Einstellung von Piloten. Ob das auf den jeweiligen Bedarf übertragen sinnvoll ist, muss bedacht werden, deshalb gibt es Verfahren, die solche Kompensationsmöglichkeiten begrenzen oder ganz ausschalten.

Das Regressionsmodell ist nur ein Modell, liegen keine quantitativen Kriterien und (wenigstens binäre) quantitativen Prädiktoren vor, eignet sich z.B.

bei binärem Kriterium die logistische Regression besser, für mehrstufige Kriterien die Diskriminanzanalyse.

Basisdesign zur Kontrastierung klinischer und statistischer Datenkombination:

Beurteiler sagen das Kriterium aufgrund von erhobenen Daten vorher. Diese Vorhersage wird verglichen mit der statistischen Vorhersage mit derselben Datengrundlage. Daraus resultieren zwei Korrelationen:

1. Statistische Kombination mit dem Kriterium r_{sk}
2. Klinischer Kombination mit dem Kriterium r_{kk}

Sind beide Vorhersageleistungen ungefähr gleich, ist die Validierung kein Mittel um eine Methode auszuwählen. Dann werden eher Kosten und Nutzen eine Rolle spielen.

Empirischen Prüfungen zufolge, wurde dieses Design nicht immer in reiner Form realisiert. Die klinische Seite erhielt den Vorzug, wobei dieser auch mehr Daten zur Verfügung standen als der statistischen, z.B. aus Interviews.

Die Studien des Psychologen Sarbin (1941, -42, -44) zur Vorhersage des Studienerfolges aufgrund von Schulabschlussnoten einerseits und Studienberater einschätzung andererseits, wurden unter der Vorannahme durchgeführt, dass klinisch diagnostische Urteile, einfach aufgrund der Kompetenz der Diagnostiker, bessere Vorhersagen erlaubten als die statistischen. Diese Erwartung erfüllte sich nicht. Korrelationen:

| Vorhersage | Männer | Frauen |
|-------------|--------|--------|
| Klinisch | .35 | .69 |
| Statistisch | .45 | .70 |

Das Ergebnis favorisiert nun nicht automatisch die statistische Seite, es läge näher zu behaupten, dass Diagnostiker zumindest empirisch informiert sein sollte. Es geht nicht um Kontroverse, sondern um Verbesserung. Auch Studien lassen mehr Fragen offen als sie beantworten können.

Die Untersuchung von Goldberg (1965) wurde sehr bekannt und griff all diese Punkte wieder auf.

184 männliche Patienten wurden untersucht. Deren Diagnosen „Psychose“ und „Neurose“ dienten als Kriterium, das statistisch und klinisch bestimmt werden sollte. Grundlage waren 11 Skalen des MMPI (Hathaway & McKinley, 1943).

Die Daten aus diesem beliebten Testverfahren zu klinisch relevanten Persönlichkeitsmerkmalen, wurden 29 unabhängigen Beobachtern vorgelegt, von denen 13 klinische Psychologen mit Dokortitel waren, die anderen Psychologen in Ausbildung. Die Beurteiler sollten die Profile auf einer elfstufigen Skala von neurotisch bis psychotisch sortieren.

Diese Daten wurden später mit dem Kriterium der psychiatrischen Diagnose verglichen.

Um zwischen Neurotikern und Psychotikern unterscheiden zu können, bildete Goldberg einen Index, den er aus fünf MMPI Skalen gewonnen hatte und an einer Stichprobe von 402 klinischen Fällen überprüft hatte, welche dort sowohl einzeln als auch kombiniert die stärksten Gewichte erhielten, also am besten diskriminierten. Hohe, positive Werte waren das Index für Psychose, niedrige waren mit der Diagnose Neurose assoziiert.

Validitätskoeffizienten (Korrelationen) und Trefferquoten (Prozentwerte, relative Häufigkeit korrekter Zuordnungen von Profilen und Diagnosen) klinischer und statistischer Datenkombinationen

| Datenkombination | Validität | Treffer A | Treffer B |
|------------------|-----------|-----------|-----------|
| Klinisch | | | |
| Spannweite; | .14 - .39 | .55 - .67 | .60 - .73 |
| Durchschnitt | .28 | .62 | .66 |
| Statistisch | .44 | .70 | .74 |

Bemerkenswerter Weise waren die Leistungen der erfahrenen klinischen Psychologen nicht besser als die der unerfahrenen Anfängerkollegen, wes-

halb die Aussagen beider zusammengefasst wurden. Die Trefferquote stieg an, als die schwer oder nicht zu entscheidenden Fälle ausschieden, aber nicht in dem Maße wie es erwartet wurde. Die einfache Linearkombination der fünf MMPI Skalen war den gemittelten Leistungen der klinischen Psychologen auch danach überlegen.

Der (eine) beste Kliniker erreichte eine Trefferquote von 73%, was im Vergleich zur statistischen Kombination ein Patt war.

Diese zwei und weitere Studien folgten durch Meehl angeregt.

Bei insgesamt 51 Studien zeigten 33 eine Überlegenheit der statistischen Prognosen, in 17 erwiesen sich statistische und klinische als gleichwertig. Die einzige Studie die auf eine Überlegenheit der klinischen Vorhersage hindeutete, wurde aus methodischen Gesichtspunkten nachträglich disqualifiziert.

Es wurden auch unterschiedliche Designvariablen untersucht, wie Veröffentlichungsdatum der Studie, Stichprobengröße, Art des Kriteriums und der Prädiktoren, Ausbildung der Beurteiler, deren Berufserfahrungen, die Informationsmenge usw.

Keine der Variablen hatte markanten Einfluss auf das Ausmaß der Überlegenheit einer Methode. Bei den acht Studien, in denen die klinische Vorhersage überlegen war, stand den Klinikern deutlich mehr Information zur Verfügung als in die statistische Vorhersage einging.

Gründe für die Überlegenheit der statistischen Datenkombination

1. **Menschliche Beurteilung**, in der sich **Alltagverzerrung und -fehler** einschleichen. Psychologisch diagnostisches Wissen und Training macht nicht immun gegen typisch menschliche Informationsverarbeitung.
2. **Inkonsistenz bei der Verwendung diagnostischer Entscheidungsstrategien**. Diagnostiker setzen bei der Urteilsbildung Entscheidungsregeln ein, verfolgen sie aber implizit und nicht konsequent, sondern wandeln sie von Fall zu Fall ab. Eine Reihe von Untersuchungen belegen, dass rigide Regelanwendung durch Diagnostiker zu besseren Entscheidungen führen als flexible.

3. [Die Möglichkeit aus Rückmeldungen](#) zu lernen ist bei der Vorhersage von Diagnostikern oft [nicht gegeben](#). Z.B. bei Berufseignung. Die Ablehnung potentieller Bewerber lässt sich im Nachhinein auf die Güte der Entscheidung hin nicht mehr überprüfen.

[Kritik, z.B. von Allport \(1942\)](#)

„Statistische Vorhersagemodelle sind irreführend, wenn sie auf einzelne Fälle anstatt auf Populationen angewendet werden. So folgt z.B. aus der Tatsache, dass 80% der Delinquenten, die aus zerrütteten Familien stammen rückfällig werden, nicht, dass dieser konkrete Delinquent, der aus einer zerrütteten Familie stammt, ein 80% Risiko hat, rückfällig zu werden. [...] ob er rückfällig wird können wir exakt vorhersagen, wenn wir alle verursachenden Bedingungen und Umstände im konkreten Fall kennen würden [...] Sein Rückfallrisiko wird durch das Muster seiner Lebensumstände determiniert, nicht durch relative Häufigkeiten in der Population. Tatsächlich ist psychologische Verursachung immer personell, niemals statistisch.“

[Bewertung des Zitats](#)

Es gibt, anstelle von Nur-Gruppendaten auch Regelmäßigkeiten, die einzelne Fälle betreffen und relative Häufigkeiten determinieren ein Rückfallrisiko natürlich nicht. Vermischt werden hier auch Häufigkeit und Ursache. Ursachen werden durch statistische Modelle nicht erfasst. Vorhersagen implizieren aber nicht zwangsläufig Verursachungen, z.B. sind zerrüttet Familienverhältnisse beim Rückfallrisiko häufiger anzutreffen, ob sie aber als Ursache anzusehen sind oder nur Mit-Ursache, wird dadurch nicht bestimmt.

Dagegen hat der Kliniker zusätzliche Informationen über die Einzelperson (Verwandte, Freunde, Erfahrung mit anderen Fällen dieser Art), die in eine statistische Vorhersage nicht eingehen.

Die Statistiker halten dagegen, dass es der empirischen Forschung überlassen sein sollte zu ergründen, welche Variablen zur Vorhersage überhaupt herangezogen werden sollten und welche unbrauchbar sind.

Vieles spricht für die breite Anwendung formeller Daten, allerdings steht ihr die mangelnde Verfügbarkeit geeigneter Modelle gegenüber. Auch die Anwendung solcher Modelle, falls verfügbar, auf jeweilige lokale Gegebenheiten ist fraglich (z.B. nicht kultur- und länderübergreifend anzuwenden). Modelle müssen also geprüft, adaptiert, modifiziert werden. Will man aber wissenschaftlich arbeiten, ist ohnehin eine regelmäßige Überprüfung des Funktionierens eines diagnostischen Systems, bei aller Kostenüberlegung, notwendig, gleichgültig um welche Seite der Urteilsbildung bei der Vorhersage es sich handelt.

Da der Disput über beide Formen der Vorhersage immer noch aktuell ist, mündete der [Versuch die Urteilsbildung auf klinischer Seite transparenter zu gestalten, in der Entwicklung paramorpher Modelle](#).

[Paramorphe Modell des Diagnostizierens](#)

Meehl wendete eine Unterscheidung wissenschaftlicher Theorien auf den Prozess klinischer Vorhersagen an, die von dem Philosophen Hans Reichenbach (1938) in folgender Form aufgestellt wurden

1. [Kontext der Entdeckung](#) = Wie kam die wissenschaftliche Hypothese zustande? Wie kam der Forscher XY dazu seine Theorie zu formulieren?
2. [Kontext der Rechtfertigung](#) = Logische Analyse der Bewährung einer Theorie

Für Meehl bezieht die der Kontext der Rechtfertigung auf die Vorgehensweise bei der Überprüfung der Gültigkeit einer Diagnose, bzw. Prognose. Die Vorgehensweise gelte für statistische wie klinische Vorhersage. Die Überlegungen zum Vergleich formeller und informeller Methoden gehören in den Kontext der Rechtfertigung.

Im Kontext der Entdeckung geht es nicht um die Validität eines Urteils, sondern darum auf welchem Wege es zustande kam (wie kommt der Diagnostiker zu seinem Urteil?)

1. Man kann einen Diagnostiker einfach fragen
2. Er kann Angaben zu seiner Regel machen nach der er vorgeht
3. Man kann ihn während der Beratung laut denken lassen
4. Man kann versuchen die Input-Output-Relation zwischen Daten und Diagnose mittels formeller Methoden zu beschreiben

Das alles führt zu einem deskriptiven (beschreibenden) Modell des Diagnostizierens. Darin ist der Diagnostiker selbst das Objekt der psychologischen Modellbildung.

Hoffmann (1960, 1968) nennt formelle Repräsentationen des diagnostischen Entscheidungsverhaltens **paramorphe Modelle**.

Dabei geht es nicht um kognitive Prozesse, sondern um den Zusammenhang zwischen der Information, die ein Diagnostiker erhebt/die ihm zur Verfügung steht und seinen Entscheidungen.

Ein bestimmtes paramorphes Modell stellt immer nur eine von mehreren möglichen Repräsentationen möglichen Entscheidungsverhaltens dar, idealer Weise ist es ein für den Fall optimales Modell. Es wird also versucht durch ein Modell darzustellen, welche Variablen ein Diagnostiker für sein Urteil als relevant erachtet.

Bei der Erstellung paramorpher Modell sind die Diagnostiker die Probanden. Sie bekommen Daten von Klienten und sollen auf der Grundlage dieser Informationen für jede Person eine spezifische Diagnose hinsichtlich eines Merkmals abgeben. Oft stehen ihm Testdaten zur Verfügung (aus Intelligenztests, Persönlichkeitstests usw.) sowie Angaben, die potentiell relevant sein können (Geschlecht, Alter). Die Daten sollten möglichst repräsentativ sein und möglichst in Zahlenform vorliegen, sowie auch das Urteil des Diagnostikers in Zahlen, z.B. auf Ratingskalen ausgedrückt werden sollte. Auf der Grundlage solche Zahlenwerte lässt sich z.B. eine Regression durchführen, um zu ermitteln, welches paramorphe Modell die Datenkombination eines Diagnostikers am besten repräsentiert.

Es gibt hierzu **lineare Modelle**. Aber auch **Konfigurationsmodelle**, in denen von den Angaben des Diagnostikers über sein Vorgehen Gebrauch gemacht wird (siehe Pilot und Raumverstehen vs. Matheverstehen, da muss geprüft werden, ob gut in Mathe zu sein, ein fehlendes Raumverstehen wirklich bei

einem Piloten kompensieren sollte). In den Modellen geht es nicht um die Bestimmung der Richtigkeit einer Diagnose, es soll ein Algorithmus erstellt werden (am einfachsten eine Gleichung), der es ermöglicht, die Diagnose bei gegebenen Daten zu wiederholen. Ein gutes paramorphes Modell liefert bei gleichen Daten die gleiche Vorhersage, wie der Diagnostiker; ob diese valide sind oder nicht ist eine andere Frage.

Lineare Modelle

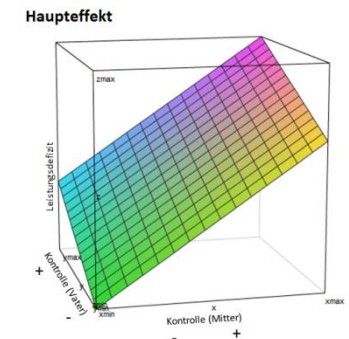
- a) Haupteffektmodelle (Haupteffekte)
- b) Modelle die kurvenlineare Beziehungen beinhalten (Quadrate)
- c) Modell mit Prädiktorinteraktion (Produkte)

(Haupteffekt und Interaktion stammen aus der Varianzanalyse)

a) **Haupteffekt** zur Vorhersage einer Diagnose, die Grundschulleistung erfasst mit den Prädiktoren „Laissez faire“ bis extreme Einschränkung und Kontrolle.

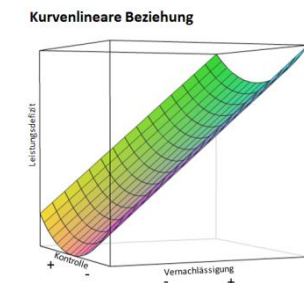
$$\hat{Y} = 0.6X_1 + 0.3X_2$$

Hier wird das Kontrollverhalten von Mutter und Vater betrachtet, wobei das der Mutter vom Diagnostiker als gewichtiger bewertet wird. Je höher die Kontrolle, desto höher schätzt der Diagnostiker das Risiko eines Leistungsdefizits ein.



b) Kurvenlineare Beziehung

Ein anderer Diagnostiker findet, dass sowohl starke Kontrolle als auch extreme Vernachlässigung für schulische Probleme verantwortlich sind. Für ihn wäre ein mittleres Maß an Kontrolle optimal. In „seinem“ Modell würde deshalb eine kurvenlineare Beziehung zwischen Kontrolle und Leistungsdefizit be-



stehen, also in der Mitte der Kurve ist das Defizit gering, an Anfang und Ende dagegen hoch.

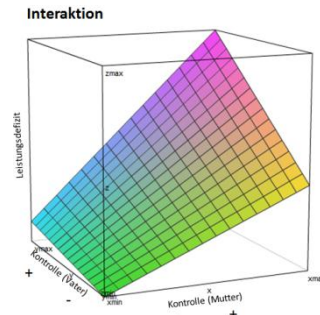
$\hat{Y} = 5 - 2X_1 + 0.2X_1^2$ ergibt eine Kurvenlinearität (U-förmig oder auch umgekehrt U-förmig).

c) Interaktion

Ein dritter Diagnostiker findet, dass die väterliche Kontrolle einen geringeren Effekt hat, wenn die mütterliche auch gering ist, dagegen habe sie einen deutlichen Effekt, wenn die mütterliche hoch sei. Hier wird der Haupteffekt (die beiden ersten Terme) durch das Produkt der Erziehungsvariablen ergänzt (gewichtete Produkte). Der dritte Term repräsentiert hier die Interaktionen zwischen den beiden Erziehungsvariablen.

$$\hat{Y} = 0.4X_1 + 0.1X_2 + 0.2X_1X_2$$

Durch Einführung weiterer Variablen lassen sich komplexere Modelle erstellen, mit der sich eine Vielzahl diagnostisch relevanter Prädiktorkombinationen abdecken lassen.



Empirische Befunde zeigen

1. Die Urteilsbildung von Diagnostikern lässt sich in vielen Bereichen bereits durch das Haupteffektmodell gut darstellen.
2. Nach Aussagen der Diagnostiker selbst, sind die „eigenen“ Modelle weitaus umfangreicher, die Befunde zeigen aber, dass sie nur glauben ihr Vorgehen bei der Urteilsbildung sei komplexer.
3. Außerdem scheinen sie die Gewichtung anders einzuschätzen als „ihr“ paramorphes Modell es dann tatsächlich abbildet.
4. Umgekehrt halten sie mehr Variablen für relevant als tatsächlich zur Modellierung ihres Urteilsverhaltens benötigt werden.

Aber die mathematischen Modelle sind auf Sparsamkeit getrimmt (möglichst wenig Prädiktor für ein wesentliches Kriterium, Indikatoren die redundant erscheinen werden rausgeschmissen)

Diese Befunde (1 – 4) sind unter der Bezeichnung „Goldbergs Paradox“ in die Literatur eingegangen. Goldberg hatte 1970 gefunden, dass die paramorphen Modelle besser diagnostizieren (Vorliegen einer Psychose vs. Neurose) als der Kliniker selbst.

Gründe

1. Diagnostiker verändern ihre Gewichtung offenbar von Fall zu Fall.
2. Sie wenden sie also in inkonsistenter Weise an.
3. Viele Diagnostiker können ihre Entscheidungsregeln nicht vollständig darstellen.
4. Sie können sie nicht verbal darstellen, weil sie ihnen nicht explizit bewusst zu sein scheinen.
5. Wottawa u.a. (1982) bemerkten, dass Diagnostiker auf Nachfragen eher Lehrbuchwissen angaben, das mit ihrem tatsächlich Verhalten aber wenig zu tun habe.
6. Obgleich Modelle eine gute Ergänzung zum Entscheidungsverhalten darstellen, werden sie von Diagnostikern oft als artifiziell empfunden und nicht als angemessene Beschreibung ihrer Urteilsbildung akzeptiert.
7. Tatsächlich lässt sich intuitiv keine Regressionsanalyse durchführen.
8. Für einen Diagnostiker sind Ausprägungsmuster von Variablen für ihre Entscheidungsfindung maßgeblich.

Die Bedeutung der Variablen hängt also davon ab, welche Ausprägungen auf anderen Variablen sie haben. Das würde sich in einem paramorphen Modell z.B. in einer Interaktion niederschlagen, die sind aber, wie kurvenlineare Beziehungen auch, schwer aufzudecken, weshalb die Modelle natürlich weniger akzeptabel erscheinen. Wenn nämlich Interaktionen nicht aufgedeckt werden, dann werden auch Kompensationen zugelassen, die in der Praxis manchmal nicht sinnvoll sind (siehe Pilotenbeispiel). Auch leidet die Akzeptanz in der Praxis daran, dass Entscheidungen über Prognosen eher

sequenziell organisierte Prozesse darstellen. Hier sind Konfigurationsmodelle geeigneter, da sie Kompensationsmöglichkeiten nur dort zulassen, wo sie sachlich gerechtfertigt erscheinen und einzelne Entscheidungsschritte, die zu einer Diagnose führen, logisch abbilden können.

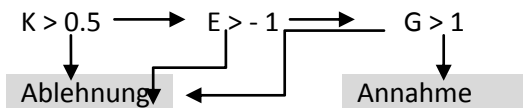
Konfigurationsmodelle

1. Sie lassen sich als Flussdiagramme darstellen.
2. Entscheidungen werden als Wenn-Dann-Regeln rekonstruiert

Das Modell heißt deshalb konfigural, weil bestimmte Konfigurationen vorliegen müssen, also bestimmte Muster von Werten bei Personen, damit eine bestimmte Diagnose gestellt werden kann, wie z.B. geeigneter Bewerber, depressiv, schulängstlich usw.)

Zwischen den Variablen auf denen die Diagnose basiert können kompensatorische Beziehungen bestehen, die ergeben sich aber nicht wie beim linearen Modell zwangsläufig.

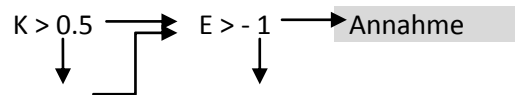
Konfigurationsmodell ohne Kompensation



Die Werte sind z-transformiert

Wenn K = Konzentrationsvermögen mindestens leicht überdurchschnittlich ausfällt **und** E = emotionale Stabilität nicht deutlich unter dem Durchschnitt liegt **und** G = Gewissenhaftigkeit überdurchschnittlich ist, wird die Annahme empfohlen, ansonsten die Ablehnung (z.B. eines potentiellen Bewerbers).

Konfigurationsmodell mit Kompensation



$V > 1 \longrightarrow$ Ablehnung

Hier führen niedriger Werte bei der Variablen K, nicht zur Ablehnung, sondern können durch hohe Werte bei $V = \text{Vigilanz}(\text{test})^1$ kompensiert werden. Kompensationen entsprechen logisch oder-Verknüpfungen von Aussagen. Ein K-Wert muss also größer als 0.5 sein **oder** der V-Wert größer als 1, damit der Bewerber nicht bereits ohne Ansehen des E-Wertes abgelehnt wird.

Erstellung und Prüfung dieser Modelle ist aufwändiger und komplexer als Regressionsmodelle, auch menschliche Supervision und Intervention sind hier stärker gefordert. Als Beispiel wird hier die Hypothesenagglutinerung angeführt:

Eine Form der Rekonstruktion diagnostischer Urteilsbildung in Form von konfiguralen Modellen wurde von Wottawa u.a. unter der Bezeichnung Hypothesenagglutinerung vorgestellt (Abk. HYPAG). Das ist ein Verfahren, das schrittweise diagnostische Entscheidungen rekonstruiert und zwar in Form logisch miteinander verknüpfter (agglutinierter) Hypothesen (Wenn-Dann-Regeln). Die Rekonstruktion der Entscheidungsregeln erfolgt in einem Dialog zwischen einem Interviewer und einem Diagnostiker. Dem Diagnostiker werden dokumentierte Fälle vorgelegt und er entscheidet z.B. im Rahmen einer Bewerberselektion, welche Bewerber er weshalb für ungeeignet oder geeignet hält. Der Interviewer formuliert über die vom Diagnostiker begründeten Entscheidungen Regeln, z.B. räumliches Vorstellungsvermögen muss einen C-Wert (Standardnorm $M = 5$ und $SD = 2$) von mindestens 4 haben, damit er im Verfahren bleiben kann.

In den nächsten Schritten kommen Stück für Stück Regeln zusammen, die am Ende ein gesamtes Entscheidungsbild ergeben. Ist das Modell mit genügend Regeln versehen, wird es auf die jeweiligen bearbeiteten Entscheidungsfälle angewendet. Kommt das Modell zu den gleichen Entscheidungen wie ein Diagnostiker, liegt ein Treffer des Modells vor, dann kann man

¹ lat., vigilis – wachsam, vigilantia – Wachsamkeit **Vigilanz** ist der medizinische Begriff für eine Bewusstseinslage (Wachheit)

zum nächsten Fall übergehen. Liegt ein Fehler vor, ist das Modell noch nicht perfekt und muss vielleicht modifiziert werden. Es könnte aber auch am inkonsistenten Vorgehen des Diagnostikers liegen oder das Modell kann noch keine Entscheidung treffen, dann greifen seine Wenn-Dann-Teile nicht. In allen drei Fällen wird der „Fehler“ mit dem Diagnostiker besprochen und behoben.

Das modifizierte Modell wird dann noch einmal auf die Entscheidungsfälle angewendet. Eventuell muss das Modell so lange iteriert werden, bis eine akzeptable Passung zwischen Modell- und Diagnostikerentscheidung hergestellt ist. Diese Passung wird vorher z.B. bei 95% festgelegt.

Danach wird eine Kreuzvalidierung an neuen, bisher zurückgehaltenen Fällen vorgenommen. Die Trefferquote bei diesen Fällen liefert einen Indikator für die tatsächliche Leistungsfähigkeit des Modells.

Zur Erinnerung: Das sagt nichts über die Validität der Diagnose aus, sondern nur über die Güte der Passung zwischen Vorhersagekraft des Modells und des Diagnostikers.

Für die Erstellung eines Modells müssen ca. 200 bis 300 Fälle bearbeitet werden. Hypothesenagglutinierte Modelle schneiden bei der Kreuzvalidierung oft besser als mit 90% Passung ab und damit besser als die linearen Modelle. Die Akzeptanz dieser Modelle ist in der Personalberatung und –auslese deutlich höher als bei klinischen Diagnostikern. Letztere sind teilweise überrascht, dass sich ihre Entscheidungsprozesse überhaupt modellieren lassen und begegnen den Modellen, obwohl es „ihre eigenen“ sind, mit Skepsis.

Vorteile expliziter Modelle

- Transparenz. Entscheidungen werden nachvollziehbar
- Optimierung. Entscheidungen werden nicht intuitiv oder von Fall zu Fall anders entschieden
- Lernen und Erfahrungsaustausch. Diagnostische Praxis ist leichter zu erlernen (Lernen am Modell)
- Entlastung. Redundanzen bei Indikatoren der Entscheidungsfindung werden vermieden

- Konsistenz. Konsistente Entscheidungen sind im Sinne des Gütekriteriums objektiv
- Validität. Paramorphe Modell bilden Entscheidungen ebenso gut ab, wie die informellen Urteile der Diagnostiker. Sie bilden ja deren Verhalten ab. Dadurch ist es möglich eine Verbesserung künftiger Entscheidungen zu bewirken.

Nutzung der klinischen Inferenz

Wesentliche Datenquelle informeller Urteilsbildung sind verbale und non-verbale Äußerungen von Klienten, im Rahmen von Interviews und anderen mehr oder weniger vorstrukturierten Erhebungssituationen erhoben.

Meehl (später fand Sawyer 1966 Belege dafür) vermutet, dass Daten aus Interviews wertvolle Beiträge zur Vorhersage darstellen, wenn sie in geeigneter Weise verwertet und aufbereitet werden, z.B. durch Erhebung und Aufarbeitung in testähnlichen Formaten.

Dieser Gedanke (formelle und informelle Modelle zusammenzuführen) wurde in jüngster Zeit weitergeführt. Entscheidend ist es die Beobachtungen aufzuzeichnen, um den Nutzen klinischer Expertise zu maximieren. Shedler & Westen entwickelten 1998 hierfür ein Q-Sort-Verfahren² zur Persönlichkeitsbeschreibung, das speziell auf den Einsatz durch klinische Experten abgestimmt ist, Shedler-Westen-Assessment-Procedure, SWAP-200.

In der neuen Version des Verfahrens zur Klassifizierung von Persönlichkeitsstörungen bei Heranwachsenden, dem SWAP-200-A sortiert ein klinischer Psychologe auf der Basis eines Interviews oder mehrerer Beobachtungen, z.B. die während einer Therapie anfallen, 200 persönlichkeitsbeschreibende Feststellungen in insgesamt acht Kategorien (0 = trifft nicht zu, bis, 7 = trifft sehr zu). Die Feststellungen beschreiben relativ verhaltensnahe Sachverhalte (z.B. rennt von zu Hause weg, oder auch manche, die ein gewisses Maß an Inferenz³ erfordern wie z.B. glaubt, dass seine Probleme durch externe

² Persönlichkeitsklassifikation. Dieses Verfahren nutzt persönlichkeitsbeschreibende Adjektive oder Feststellungen, die in vorgegebene Kategorien sortiert werden. Die Kategorien stellen das Ausmaß dar, in dem das Adjektiv auf die zu beurteilende Person zutrifft.

³ Schlussfolgerung

Faktoren bedingt sind). Alle Items sind relativ einfach und ohne Terminologie einer spezifischen therapeutischen Schule formuliert, damit eine breite Anwendung erfolgen kann. Die verwendeten Feststellungen wurden in Zusammenarbeit mit klinischen Psychologen auf ihre Brauchbarkeit hin überprüft. Das Verfahren wurde außerdem mit den gängigen psychometrischen Methoden optimiert.

Westen et.al. (2003) sammelten mit diesem Verfahren professionelle Beurteilungen von mehreren hundert Heranwachsenden (Adoleszenten), die in Behandlung waren. Mittels eine statistischen Klassifikation der von den Psychologen und Psychiatern vorgenommenen Q-Sortierungen konnten sie fünf prototypische Muster, also distinkte Gruppenprofile identifizieren

1. Antisoziale-psychopathisch
2. Emotional dysreguliert
3. Vermeidend
4. Narzisstisch
5. Gehemmt-selbstkritisch

Bei den ersten vier handelt es sich um klinisch relevante Formen von Persönlichkeitsstörungen, bei der letzten eher um einen Persönlichkeitsstil. Nummer 1 z.B. umfasst folgende Beschreibung:

Rebellisch, aufsässig gegen Autoritätsfiguren, intensiver und situationsunangemessener Ärgerausdruck, Schuld für die eigenen Fehler bei anderen suchen, Wut und Erniedrigung als Reaktion auf Kritik, Unzuverlässigkeit, mangelnde Bereitschaft, Verantwortung zu übernehmen. Die statistisch bestimmte Zugehörigkeit zu dieser Gruppe war mit einer Reihe externer Kriterien assoziiert. Hierher gehörten z.B. schlechte schulische Leistungen, Drogenmissbrauch, Kriminalität und Erfahrung physischer Gewalt in der Kindheit.

Datenerhebung durch Beobachtungen, Gespräche, auch mit Bekannten Familienmitgliedern, Freunden, sowie Inferenzen und Integration von Daten obliegen dem klinischen Psychologen. Den Klinikern wird aber keine Vorhersage abverlangt z.B. derart, ob eine Person in den nächsten 6 Monaten mit dem Gesetz in Konflikt gerät, auf eine Therapieform anspricht oder

erfolgreich im Studium sein wird. Diese Aussagen sollen auf der Basis formeller Datenkombinationen erbracht werden können. Klinische Inferenz und formelle Klassifikation lassen sich auf Datenebenen, die standardisierte Beschreibungen liefern miteinander verbinden und so können besser Prognosen gestellt werden.

Eine Konsequenz aus dieser Erkenntnis bestand in dem Bemühen einheitlich detaillierte Richtlinien zur Bestimmung von Störungen zu entwickeln, die sich in einem internationalen Klassifikationssystem niederschlugen.

| Klinische Tätigkeit | Formelle Prozedur |
|--|-------------------|
| Datensammlung Inferenz und Kombination | Klassifikation |
| Q-Sortierung oder andere Formen systematischer Beschreibungen | Vorhersage |

Petermann spricht, als Konsequenz aus geschilderten Diskussion und der Erkenntnis, dass die Verwendung expliziter Regeln, die Kontrolle über Richtigkeit und Angemessenheit von Diagnosen und Prognosen für die Verbesserung der Entscheidungsfindung sorgt, von „kontrollierter Praxis“.

Pflichtliteratur, Krohne und Hock (2007), Psychologische Diagnostik und Anwendungsfelder. Kapitel 8: Handlungstheoretische Modelle

8.1 Grundbegriffe der Handlungstheorie

Strategien sind an die Entwicklung von Intentionen und Zielen gebunden. Formulierung von Zentrale Begriffe der Handlungstheorie

- Strategie
- Intention

- Ziel

Handlung ist in der Psychologie eine **zeitlich organisierte, zielgerichtete Tätigkeit**.

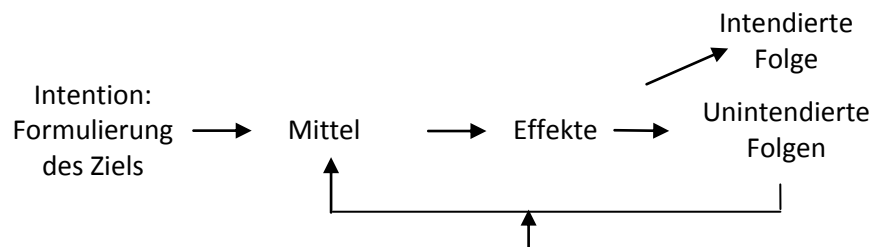
Verhalten dagegen kann zielgerichtete Aktivität sein, aber auch das reine Reagieren auf externe und interne Stimuli.

Vor der Handlung muss das Ziel **antizipiert** sein, also die kognitive Repräsentation muss als Antizipation⁴ gespeichert sein und während des Handelns zugänglich gehalten werden. Handlung fordert auch Einsatz von bestimmten **Mitteln**, deshalb muss zuerst **geplant und organisiert** werden, weshalb diese drei gleich unter dem Begriff Tätigkeit zusammengefasst werden.

Der Einsatz spezifischer Mittel führt zu bestimmten **Effekten**, die wahrgenommen werden müssen und im Hinblick auf die antizipierten Ziele beurteilt. Das gilt auch bei formulierten Zwischenzielen. Die Wahrnehmung und Beurteilung des Effekts bewirkt, dass die Folgen des eigenen Tätigseins realisiert werden. Wurde das **Ziel erreicht, ist die Folge intendiert**.

Die eigene Tätigkeit führt aber auch zu **unintendierten Folgemengen**, die weder beabsichtigt noch antizipiert waren.

Die Wahrnehmung nicht beabsichtigter Folgen eingesetzter Mittel, führt zu einem **weiteren Beurteilungsprozess (Reflexion)** beim Handelnden. Inhalt der Reflexion ist die **Abschätzung der Tolerierbarkeit solcher Folgen**. Sollten nach Einschätzung des Reflektierenden nicht alle Folgen tolerierbar sein, wird eine neue Handlung in Gang gesetzt. Das Ziel dieser Handlung ist die **Reduzierung des Wirksamwerdens dieser Folgen (Schadensbegrenzung)**.



⁴ Vorwegnahme

Reflexion des Handelnden

8.1 Handlungstheorie und Psychodiagnostik

Exkurs: Diagnostizieren ist kein Entdeckungsprozess. Es geht nicht darum bei einem Klienten alles oder etwas aufzudecken, in Sinne von „wie ist der Klient“.

Moderne Diagnostik ist ein Prozess des Planens, Veränderens, Entscheidens und Beurteilens. Es geht darum Zielsetzungen zu realisieren.

In diesem Kapitel geht es um die Bedeutung des Diagnostizierens bezogen auf den Einzelfall und die Beschreibung des engeren Prozesses der diagnostischen Urteilsbildung und den gesamten Ablauf interventionsbezogener diagnostischer Arbeit.

Nach Kaminski umfasst dieser Prozess kognitive Mikrokomponenten der Beurteilung und Entscheidung, der auf der Aktivierung spezifischer Wissensrepräsentationen basiert. In diesem Kapitel werden besprochen:

1. Modell des Arbeitsflusses in der Diagnosepraxis
2. Implikationen des Modells
3. Konkretisierung des Modells am Beispiel
4. Analyse spezifischer Aspekte

8.2.2 Modell des Arbeitsflusses

Bei der diagnostischen Urteilsbildung handelt es sich um einen sequentiellen Arbeitsprozess mit Rückmeldungsschleife, der erst abgeschlossen ist, wenn das zuvor definierte Ziel erreicht ist.

Der Prozess ist eine Verflechtung aus Auftrag, Diagnose und praktischer, z.B. modifikatorischer Intervention.

1. Phase: Eingangsdaten

Die kommen von selbst: Klagen, Wünsche, Fragestellungen von außen. Grundlage des Auftrags, aber noch nicht Auftrag

2. Phase: Hypothesenbildung und –bearbeitung

Psychologische Beschreibung des Eingangszustandes, Ursachen des Zustandes und Definition des Zustandes auf den hin geändert werden soll.

Beispiel: Schulunlust. Die Ursache könnte als Vermeiderverhalten beschrieben werden. Ursachen: Misserfolg als Konsequenz fehlender Kompetenz oder z.B. eine erhöhte Furcht vor Misserfolg.

Zielzustand: Verringerung der Schulunlust, Erhöhung der Kompetenzen. Der Weg dorthin könnte im Aufbau fachlicher Kompetenzen liegen, oder eben in einer Modifikation zur Beseitigung von Misserfolgserwartungen durch „Umattribuieren“. Denkbar wäre auch eine Modifikation des hohen elterlichen Leistungsanspruches.

Psychologisches Grundlagenwissen spielt eine große Rolle bei der diagnostisch-praktischen Urteilsbildung.

3. Phase: Hypothesenbeurteilung

Kann man die Fragen der 2. Phase „wie ist der gegenwärtige Zustand des Klienten und auf welchen Zustand soll hingearbeitet werden, wie kommt man vor allem dahin?“ nicht präzise beantworten, was in der Praxis häufig vorkommt, nutzt man die Beurteilungsphase um zu mehr Informationen zu gelangen, man kann sonst keine Hypothese formulieren.

4. Phase: Planung der Datenbeschaffung

Zu wenig Daten, keine Hypothese, deshalb Beschaffung von mehr Daten.

Beispiel zur Schulunlust:

Hypothese 1: Häusliche Spannungen, Tod einer Bezugsperson. In der Theorie muss gesagt werden, wie sich diese Ereignisse auf die Schulleistungen auswirken. Dann können diesbezüglich gezielte Fragen gestellt werden, Interaktionen zwischen Familienmitgliedern beobachtet und spezifische Tests besorgt werden.

Hypothese 2: Intellektuelle Ursachen, kann durch Intelligenz- oder Schulleistungstests geprüft werden.

Hypothese 3: Organische Ursachen. Beschaffung von internistischen oder neurologischen Daten.

5. Phase: Planung der Praxis

Sobald Wissen über das „Wie“ der Änderung vorliegt, wird das praktische Vorgehen geplant

6. Phase: Praxis

Beratung, Selektion bzw. Klassifikation, Gutachten

7. Phase: Prüfung der Praxis

Verläuft die Praxis in Übereinstimmung mit der Hypothese?

Gibt es Verbesserungen im Laufe der Intervention? Wenn nicht, muss die (Phasen-)Schleife neu durchlaufen werden. Zusätzliche Information können plötzlich dazu kommen. Wurde das Zielkriterium erreicht?

8.2.3 Implikationen des Modells

- (1) Alle diagnostischen Auswertungen von Daten haben den Charakter von Hypothesen oder auch hypothetischen Interpretationen. Sie werden erfunden, provisorisch als gültig angenommen, beurteilt, geprüft, beibehalten oder verworfen
- (2) Datenbeschaffung wird geplant und ist Hypothesen geleitet, auch das routinemäßige Testen (Intelligenztest usw.).
- (3) Prüfungs- und Entscheidungsprozesse müssen ebenso klar definiert sein, wie das Hypothesensetzen und Planen der Datenbeschaffung
- (4) Formulierung diagnostischer Hypothesen ist Grundvoraussetzung zur Einleitung der praktischen Phase.
- (5) Ziele sollten so formuliert werden, dass sie auch zu erreichen sind. Der Arbeitsprozess fliegt dann solange eine Schleife, bis das Zielkriterium erreicht wurde.

Jetzt folgt eine Beschreibung darüber, dass in der Praxis einige Phasen einfach übersprungen werden und welche Konsequenzen daraus entstehen. Auffällig war die kritische Auseinandersetzung mit der Psychoanalyse, von der gesagt wird, sie haben ein Ursachenportfolio und wende dies quasi automatisch auf festgelegte Störungsbilder an, ohne davon auszugehen, dass es auch andere Ursachen geben könnte, was die nachfolgend Intervention natürlich ebenso beeinträchtigt. Es wird kritisiert, dass ohne Formulierung und Prüfung von Hypothesen im Grunde keine diagnostische Tätigkeit möglich sei.

8.2.4 Konkretisierung des Modells

Dieses Kapitel überspringe ich, weil hier noch einmal die einzelnen Phasen expliziert werden, und das für meine Begriffe bei jedem, der die Praxis kennt, offene Türen einrennt. Nur, dass bei der Beschreibung von Selbstverständlichkeiten hochtrabendere Wörter eingesetzt werden, als die Menschen an der Peripherie sie je benutzen würden.

So wird vom Wissen darüber, auf welche Weise Menschen geändert werden können als

Speicher 1 = Änderungswissen

gesprochen und, dass der Psychologe eventuell auf Fremdkompetenz zurückgreifen muss, wird als

Speicher 2 = Kompetenzwissen

bezeichnet.

Sagen Speicher 1 und 2, dass des Psychologen psychologische Kompetenz nicht ausreicht, muss ein anderer Fachmann ran.

Kriegt man im Speicher 1 bereits heraus, dass das Problem nicht geändert werden kann (z.B. Down-Syndrom), dann beeinflusst das die Zielsetzung.

Speicher 1 = Benötigt auch Grund- und Erfahrungswissen

Speicher 3 = Erfahrungswissenschaftlich überprüfte Änderungsmodelle⁵ (des Speichers 1) werden von Kaminski nach zwei Aspekten beschrieben:

1. Benennungen für Dispositionen am Individuum, die geändert werden sollen. Beispiel: Gewohnheit, Erwartungen, Kompetenzen, Einstellungen, Werte oder Ziele
2. Benennungen für Typen von Änderungsprozessen, die sich an diesen Dispositionen abspielen. Beispiel: Üben, Extinktion⁶, Rollenübernahme, Differenzierung oder Beobachtungslernen

Dabei existiert das Problem der Wissensanwendung. Es muss natürlich geprüft werden, ob die eingehenden Daten mit dem Änderungsmodell äqui-

valent sind. Bei Schulunlust, die auf Gewohnheit basiert, ist das Änderungsmodell wirksam, z.B. mit Umkonditionieren. Grundet sich die Schulunlust auf einer familiären Konfliktsituation zwischen den Eltern, wird das Änderungsmodell anders aussehen. Eine unterschiedliche Kategorisierung hinsichtlich der Änderung muss also vom Psychologen vorgenommen werden.

Speicher 4 = „Gewissen“

Mindestanforderung hinsichtlich der Zustände an Menschen, die unbedingt verwirklicht werden müssen, bzw. deren Fortbestehen nicht toleriert werden kann. Die Frage danach stellt sich vor allem, wenn verschiedene Zielsetzungen von außen an den Psychologen herangetragen werden. Beispiel: Der Vater will, dass der Sohn auf dem Gymnasium bleibt, der Schulleiter empfiehlt eine andere Schule, wegen Überforderung des Kindes.

Zielzustände in grundwissenschaftlichen Änderungsmodellen werden in einer operationalen Sprache beschrieben. Hier ist z.B. ein bestimmtes Lernkriterium der Zielzustand, in diesem Falle ginge es um eine kriteriumsorientierte Messung. Hier müssten sowohl Wohlbefinden des Kindes als auch Leistung operationalisiert werden, durch Fremd- und Selbsteinschätzung, Beobachtung, Leistungstests u.ä. Selbstbericht und Verhaltensbeobachtung sind in ihrem Erhebungsmodus oft nicht zeitgleich, weshalb eine operationale Definition notwendig ist, um eine objektive Entscheidung darüber zu treffen, ob ein Ziel erreicht wurde oder nicht. Andernfalls kann es zur vorzeitigen Beendigung der Modifikation kommen, sowohl seitens des Klienten als auch seitens des Psychologen.

Bekannt geworden ist in diesem Zusammenhang der Hello-good-bye Effekt von Hathaway (1948). Anfang einer Therapie (hello) und Ende einer Therapie (good bye) sollte beinhalten, dass sich Klient und Therapeut einig sind, dass der Klient eine Therapie benötigt, bzw. dass es dem Klienten am Ende wieder besser geht. Die Einigkeit darüber sagt aber nichts über die Nachhaltigkeit der Besserung aus.

8.3 Die Bewertung des Modells

⁵ Ich bin hier unsicher, weil der Speicher 3 nirgends explizit erwähnt wird, der Speicher 4 aber schon, was bedeutet, es muss einen Speicher 3 irgendwo gegeben haben

⁶ Kommt aus der Optik: Auslöschung, Abschwächung

Kaminski hat mit Vorlegen seines Modells den Prozess des Diagnostizierens von einem einmaligen Erkenntnisakt hin zu einem sequentiellen prozesshaften Vorgehen mit Rückkopplung beschrieben. Für ihn ist der gesamte Arbeitsfluss des psychologischen Intervenierens, in dem das Stellen der Diagnose nur einen Teil darstellt, nach dem Vorbild von grundwissenschaftlichen Forschungsstrategien zu konzipieren. Jede Phase soll durch präzise, operationalisierbare Hypothesen gesteuert werden, wobei das Zwischenergebnis wiederum den weiteren Arbeitsablauf bestimmt. Die Anwendung seiner Reihe von Unterscheidungsdimensionen erlaubt unterschiedliche Varianten des Diagnostizierens zu klassifizieren.

Kaminskis Ziele waren:

- (1) Das Tätigkeitsfeld des Praktikers mit Hilfe grundwissenschaftlicher psychologischer Denkweisen durchschaubar zu machen
- (2) Praktiker verschiedener Teildisziplinen sprechen manchmal verschiedene Sprachen, ohne zu prüfen, ob ihre Theorien miteinander vereinbar sind. Ziel wäre hier eine konsistente Theorie, etwa die eines kognitiven Prozesses der Informationsverarbeitung und Problemlösung (Ein Ansatz der 1998, von Grawe, Donati und Bernauer unter „Allgemeine Psychotherapie“ vorgestellt wurde)
- (3) Praktiker wissen zu wenig, was sich im grundwissenschaftlichen Fundament der Diagnostik abspielt, weshalb Kaminski das Konzept der Speicher entwickelt hat.

Ansicht der Autoren

Kaminskis Überlegungen sind eigentlich keine Theorie, sondern eher eine rationale Rekonstruktion des diagnostischen Prozesses. Ihm ging es darum den Diagnostikprozess grundlegend und vollständig durchschaubar zu machen. Seine „Theorie“ ist eher eine grundwissenschaftlich fundierte Anleitung für praktisch arbeitende Psychologen. Das handlungstheoretische Schema dabei informiert den Praktiker darüber, wie er Interventionen in verschiedene Arbeitseinheiten einteilen kann, wie er sie organisiert und einen zuvor operational definierten Zielzustand optimal erreichen kann, ohne zu viele unintendierte Folgen dabei zu produzieren.

Anmerkung von mir

Mehr gab das Kapitel für meine Begriffe nicht her, es war künstlich aufgepustet und wie ich finde unnötig, dennoch hoffe ich, dass ich das Wesentliche auch herausgefunden und nichts übersehen habe.