3. Grundlagen der Messung

3.1. Definitionen der Messung

Campbell (1920): "... die Zuweisung von Zahlen, um Eigenschaften darzustellen."

Stevens (1951): "Messung ist die Zuordnung von Ziffern zu Objekten oder Ereignissen nach bestimmten Regeln."

Gutjahr (1972),u. a.: "Zuordnen von Zahlen zu Objekten, wobei die Relationen zwischen den Objekten durch die Relationen zwischen den Zahlen reflektiert werden soll."

Die Messung soll an einem Beispiel deutlicher werden:
Eine junge Frau soll aus einer Entfernung von 5 Metern einen attraktiven Mann betrachten. Sie soll den Mann bezüglich des Ausprägungsgrades von 5 Eigenschaften einschätzen:
Liebenswürdigkeit,
Charakterstärke,
Persönlichkeit,
Musikalität
und Intelligenz.
Es wird eine Skala mit den Ziffern 1 bis 5 vorgegeben, wobei 1 eine nur sehr schwache Ausprägung der betreffenden Eigenschaft bedeutet und 5 eine sehr starke Ausprägung. Sie soll also anhand der Zahlen ausdrücken, wie liebenswürdig, wie charakterlich stark usw. sie den Mann einschätzt. Dann

wird eine andere Frau gebeten, den selben Mann einzuschätzen. Die Ergebnisse der ersten werden mit denen der zweiten Frau verglichen. Danach beurteilen beide Frauen noch eine Anzahl weiterer Männer.

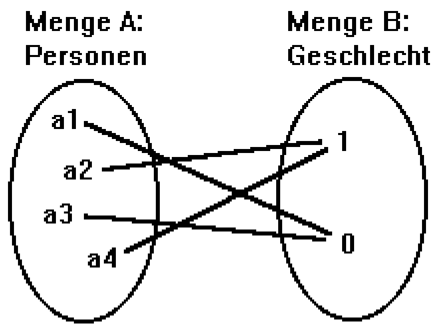
Im täglichen Leben sind wir dem gleichen Verfahren ausgesetzt wie der Mann, und wir wenden oft ähnliche Verfahren auf unsere Umwelt an. Nach den obigen Definitionen handelt es sich um ein Beispiel für Messung:
Die Ziffern 1 bis 5 wurden Personen, nämlich den Männern zugeordnet, wobei die Regel für diese Zuordnung die Anweisung zur Einschätzung einhielt.

Ziffern sind hierbei Symbole. Wenn eine Ziffer quantitative Bedeutung erlangt (Wieviel-Frage ist möglich), wird sie zur Zahl. Die Ziffer kann aber auch nur zur Bezeichung von Objekten verwendet werden, z.B.: bei Fussballspielern oder Billardkugeln.

Zuordnung heisst "in Beziehung setzten". Gegeben seien 2 Mengen. Den Elementen der einen Menge werden die Elemente der anderen Menge zugeordnet. In der Psychologie oder Erziehungswissenschaft werden meist Zahlen Personen zugeordnet.

Die Regel für die Messung ist die Anweisung was man zu tun hat. Eine Funktion ist eine mathematische Regel.
Bsp.: Weise einer sehr liebenswürdigen Personen eine 5 zu, nicht liebenswürdige Personen erhalten eine 1 oder, wie in Abbildung 1 weise männlichen Personen eine 0, weiblichen eine 1 zu.

Abbildung 1: Zuordnung von Zahlen zu Objekten



Legende: Den Personen der Menge A (a1: Person 1 ; a2: Person 2; a3: Person 3; a4: Person 4) werden Zahlen der Menge B zugeordnet (1: weiblich; 0: männlich)

Allgemein gilt: Beim Messen wird jedem Element der Menge A (dem Objekt) ein Element der Menge B (den Ziffern) durch eine Funktion zugeordnet:

f = {(a,b); a = Objekt und b = Ziffer}

Es gibt gute und schlechte Zuordnungsreglen (Funktionen). Entsprechend ist auch die Messung eine gute oder schlechte. Geschlecht ist einfach zu messen und es gibt klare Kriterien. Weniger klar sind die Kriterien im Beispiel der Beurteilung der Eigenschaften des Mannes anhand des ersten Eindrucks.

Merke: Je besser und klarer die Regeln, desto besser die Messung.

Das Problem in der Psychologie ist, dass wir psychische Eigenschaften nicht so direkt feststellen können, wie physische Eigenschaften (Grösse, Gewicht, Haarfarbe). Genau genommen können wir nur Verhalten beobachten und dieses als Indikator (oder Hinweis) auf eine Eigenschaft nehmen. Wenn ein Junge ständig andere Jungen schlägt, kann dieses Verhalten ein Indikator für zugrundeliegende Aggressivität sein.
Menschen, die sich in derselben Reizsituation befinden, können sich sehr unterschiedlich verhalten. Ein zweiter Junge steht vielleicht weinend in der Ecke. Beide Jungen haben zuvor eine schlechte Mathematiknote erhalten. Diese beobachtbaren unterschiedlichen Reaktionen bezeichnet man als Verhaltenseigenschaften. Beobachtet man das Verhalten in verschiedenen Situationen, kann man auf eine zugrundeliegende Disposition (Handlungsbereitschaft, Trait) schliessen, die aber nicht direkt beobachtbar ist, sondern ein gedanklich konstruierter Zusammenhang darstellt.
In der Psychologie verwenden wir für diese zugrundeliegenden (latenten) Eigenschaften den Namen Konstrukte. Andere Beispiel für Konstrukte wären:
Autorität, Leistung, soziale Schicht, Kreativität oder auch Intelligenz.
Um diese Konstrukte zu messen, brauchen wir Regeln. Durch die Operationalisierung (Messbarmachung) definieren wir die Regeln, wie man bei der Messung vorgehen muss. Nehmen wir als Beispiel das Konstrukt "Intelligenz". Intelligenz können wir operational definieren durch den erreichten Gesamtwert des IST70. Durch das Verhalten bzw. die Leistung beim Ausfüllen des Tests schliessen wir auf die zugrundliegende Eigenschaft.

3.2. Problembereiche des Messens

3.2.1 Wechselwirkung beim Messvorgang

Bei Messvorgängen sind zwei Systeme beteiligt, das Messinstrument (z.B. ein menschlicher Beurteiler wie im obigen Fall die Frau) und der zu messende Gegenstand (z.B. das Urteilsobjekt wie im obigen Beispiel der Mann). Der Messprozess ist die kontrollierte Wechselwirkung beider Systeme, mit dem Resultat eines Messwertes auf dem Messinstrument (z.B. die Frau sagt: "Liebenswürdigkeit erhält den Wert 5").
Wechselwirkung bedeutet nun, dass nicht nur das Messobjekt das Messinstrument beeinflusst, sondern auch dass das Messinstrument das Messobjekt verändern wird (z.B. der Mann wird sich anders verhalten, da er um die Messung weiss oder zumindest sich beobachtet fühlt). Lernen, Ermüdung, soziale Erwünschtheit, Habituation u. a. sind Dinge, die beim Messprozess berücksichtigt werden müssen.

3.2.2 Repräsentationsproblem

Für die Messung muss geklärt werden, welchen Relationen der Messwerte welche Relationen der Messobjekte entsprechen soll. Es soll einem "empirischen Relativ" (Messobjekte mit ihren Eigenschaften und Relationen untereinander) durch eine Abbildungsregel (z.B. "ist grösser als") ein "numerisches Relativ" (z.B. die Menge der reellen Zahlen **R**) zugeordnet werden (vgl. Suppes & Zinnes, 1963).
In Abbildung 1 sind die beiden Mengen A und B dargestellt. Wird jedem Objekt genau eine Zahl zugeordnet und kann man von der Zahl wieder eindeutig das Objekt identifizieren, dann sprechen wir von isomorpher (eineindeutiger) Abbildung (Bsp.: Fussballspieler). Wenn aber beispielsweise mehreren Schülern derselbe Intelligenzquotien zugeordnet werden kann, kann kein Rückschluss mehr erfolgen von der Zahl auf genau ein Objekt. In diesem Fall sprechen wir von homomorpher Abbildung, die in der Psychologie sicher üblicher ist, was wiederum Auswirkungen auf die zugrundeliegende Skala hat.
Messen kann nun definiert werden als homomorphe Abbildung oder Repräsentation eines empirischen Relativs in ein brauchbares numerisches Relativ (vgl. Kranz, Luce, Suppes & Tversky, 1971).
Das Repräsentationsproblem fragt nach den Bedingungen unter denen sich das empirische Relativ durch ein numerisches repräsentieren lässt.

3.2.3 Eindeutigkeitsproblem

Die Abbildungsfunktion bestimmt wie das empirische Relativ homomorph durch ein numerisches Relativ abgebildet wird. Es resultiert eine Skala mit bestimmten Eigenschaften. Das Eindeutigkeitsproblem fragt nach der Eindeutigkeit der Messwerte, d.h. welcher Skalentyp (s.u.) zugrundeliegt, und welche Veränderungen (Transformationen) der Messwerte möglich sind, ohne die Eigenschaften der Skala zu verändern.

3.2.4 Bedeutsamkeitsproblem

Die Bedeutsamkeit einer numerischen Aussage ist gegeben, wenn sich durch die zulässigen Transformationen die Aussage nicht ändert. Dieses Problem richtet sich auf die rechnerische Weiterverarbeitung und schliesslich die statistische Interpretation der Skalenwerte. Welche statistischen Kennwerte bei welchem Skalentyp berechnet werden dürfen wird im folgenden Abschnitt dargestellt.

3.3. Was bedeutete eigentlich Skalierung

Skalierung ist die Einrichtung einer Skala zur Durchführung einer Messung. Es lassen sich verschiedene Skalentypen unterscheiden je nachdem, wie eindeutig eine Skala ist. Die Regeln, nach denen die Ziffern den Objekten zugewiesen werden sind die Kriterien, welcher Skalentyp zugrundeliegt. Man unterscheidet hauptsächlich fünf Skalentypen:

1. Nominalskala
2. Ordinalskala
3. Intervallskala
4. Verhältnisskala
5. Absolutskala

<u>3.3.1 Nominalskala</u>

Die Zahlen, die den Objekten zugeordnet werden haben keinen numerischen Charakter, d.h., sie können nicht geordnet oder addiert werden. Die Zahlen haben hier qualitative, nicht quantitative Bedeutung. Bsp.: Numerierung von Fussballspielern, Kontonummern, Geschlechtszuweisung. Auch die willkürliche Einteilung in 2 Gruppen wie hoch-niedrig, jung-alt usw. sind Nominalmessungen. Für die Nominalmessung ist wichtig, dass man Zahlen voneinander unterscheiden kann (z.B.: 1 unterscheidet sich von 2). Veränderungen (Transformationen) der zugewiesenen Zahlen sind solange zulässig, solange man die zugewiesenen Zahlen noch unterscheiden kann (z.B.: weibliche Personen erhalten anstelle der 1 eine 526 und männliche anstelle der 0 eine 13). Die statistische Verarbeitung erfolgt durch feststellen der Häufigkeiten pro Kategorie (wieviele Objekte fallen in eine Klasse). Weitere Kennwerte wie Chi ^{2}, Kontingenzkoeffizienten und punkttetrachorische Korrelation können berechnet werden. Die Kreuztabellen "SPSS: Crosstabs" finden häufig Anwendung: Beispiel in Tabelle 2: Vier-Felder-Tafel: Lehrer mit staatsbürgerliche Motivation hoch vs. niedrig mal katholischer vs.. evangelischer Glauben.

Tabelle 2: Vier-Felder-Tafel

		Lehrer mit staatsbürgerlicher Motivation	
		niedrig	hoch
Glaubens-zugehörigkeit	katholisch	N [a] (kath./ niedrig)	N [a] (kath./ hoch)
	evangelisch	N [a] (ev./ niedrig)	N [a] (ev./ hoch)

Anmerkung: [a] N: Anzahl Personen

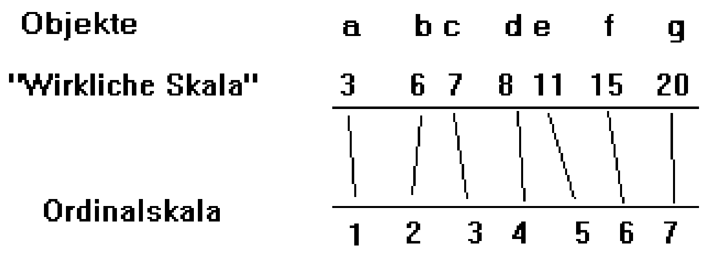
3.3.2 Ordinalskala

Die Objekte können in einer Rangreihe angeordnet werden (z. B.: die Vorlesung Statistik ist spannender als Testtheorie). Es muss gelten:

Wenn (a > b) und (b > c), dann gilt (a > c) (Transitivitätspostulat).

Die Rangzahlen bestimmen die Rangordnung der Objekte. Die Intervalle zwischen den Zahlen müssen nicht gleich gross sein. Es gibt keinen absoluten Nullpunkt (Man kann nicht sagen wann das gemessene Merkmal nicht vorhanden ist). Für das ordinale Messen ist wichtig, dass sich die zugewiesenen Zahlen unterscheiden und, dass sie sich ordnen lassen. Veränderungen (Transformationen) sind solange zulässig, solange die Ordnung erhalten bleibt. Solche Transformationen nennt man monotone Transformationen, d. h. die Beziehung zwischen den Ursprungszahlen und den transformierten Zahlen ist kontinuierlich wachsend bzw. fallend. Man könnte also eine Konstante addieren, die Ursprungszahl quadrieren oder Wurzel ziehen und anderes mehr. Bsp.: Schulnoten, Testrohwerte, Ratings, Richtersche Erdbebenskala. Werte aus Intelligenz-, Eignungs- oder Persönlichkeitstests sind häufig nur ordinal skalierbar.
In Abbildung 3 werden den Objekten a bis g die Rangzahlen 1 bis 7 zugeordnet. Wie die Objekte "in Wirklichkeit" zueinander stehen, ist bei psychologischen Merkmalen oft nicht festzustellen.

Abbildung 3: Zuordnung von Zahlen nach ihrem Rang (Ordinalskalierung)



Legende: Den Objekten a bis g mit den Ausprägungen auf der "wirklichen Skala" werden auf der Ordinalskala die Ränge 1 bis 7 zugeordnet.

Die statistische Verarbeitung erfolgt oft anhand von Median, Perzentile und Spearman Rangkorrelation.

<u>3.3.3 Intervallskala</u>

Die Abstände (Intervalle) zwischen den gemessenen Objekten sind gleich. Den Objekten werden so Zahlen zugeordnet, dass gleiche Differenzen zwischen den Zahlen gleichen Differenzen in den Ausprägungen des gemessenen Merkmals bei den Objekten entsprechen. Für die Intervall-Messung ist wichtig, dass die zugewiesenen Zahlen sich unterscheiden, dass sie eine Ordnung haben und dass die Differenzen zwischen den Zahlen einen Sinn ergeben. Veränderungen (Transformationen) können gemacht werden wenn die Verhältnisse der Abstände zwischen den Messwerten unverändert bleibt. Dies ist der Fall bei linearen Transformationen der Art:

$$X' = m * X + k$$

mit:	X'	... transformierte Zahl
	m	... konstanter Multiplikator, der nicht Null sein darf.
	X	... Ursprungszahl
	k	... additive Konstante

So kann zum Beispiel die Fahrenheit-Skala zur Messung der Temperatur in die Celsius-Skala transformiert werden durch die Formel:

$$C = \frac{5}{9} * (F - 32)$$

Bsp.: Temperatur in Celsius, Jahreszahlensystem, Standardtestnormen, Nutzenmessungen. Die statistische Verarbeitung erfolgt durch Mittelwert, Varianz und Pearson Produkt-Moment-Korrelation.

3.3.4 Verhältnisskala

Diese Skala besitzt zusätzlich noch einen absoluten Nullpunkt. Ergibt eine Messung Null, so hat das gemessene Objekt nichts von der gemessenen Eigenschaft. Für die Verhältnismessung ist wichtig, dass die zugewiesenen Zahlen sich unterscheiden, dass sie eine Ordnung haben, dass die Differenzen zwischen den Zahlen einen Sinn ergeben und dass der Messwert Null bedeutet, dass das gemessene Merkmal nicht vorhanden ist. Es sind Transformationen der Messwerte durch Multiplikation oder Division möglich. Der Nullpunkt darf allerdings nicht verändert werden (z. B.: durch Addition einer Konstanten). Wenn man zum Beispiel über eine Geldmenge von Fr 50.- verfügt, so kann man berechnen, dass dies c.a. DM 58.- (Formel: DM = 1.16 * Fr) entspricht. Der Wert Null bedeutet, dass man über kein Geld verfügt, egal in welcher Währung. Diesen Wert darf man nicht durch entsprechende Transformation verändern. Zudem kann man sagen, dass die Geldmenge von DM 100.- doppelt so gross ist, wie die von DM 50.-. Bsp.: Länge, Gewicht, Zeitintervalle, Preise.

3.3.5 Absolutskala

Zu den Eigenschaften der Verhältnisskala hinzu kommt noch, dass man bei der Absolutmessung eine natürliche Masseinheit hat. Die einzige Skala, die dies erfüllt ist die Häufigkeitsskala. Es sind keine Transformationen zulässig, da nur die absolute Häufigkeit die volle Information enthält. Bsp.: Menge von Menschen im Hörsaal, Anzahl "äh" des Dozenten pro Stunde, usw.

Tabelle 3 fasst die hauptsächlichen Skalentypen zusammen.

Tabelle 3: Kerlinger (1979, S. 671)

Skalenniveau	zulässige Aussagen	Statistiken	Beispiele
Nominal:	Gleichheit Verschiedenheit	Frequenz-statistik, Kontingenz-koeffizient, Vierfelder-korrelation	Geschlecht dichotomisierte Merkmale
Ordinal:	größer-kleiner	Median, Perzentil, Rangkorrelation	Testrohwerte, Ratings ...
Intervall:	Gleichheit von Intervallen und Unterschieden	Mittelwert, Varianz, Produkt-Moment-korrelation	Celsius, Standardtest-normen
Verhältnis:	Gleichheit von Summen, Vielfachen und Quotienten	geometrisches Mittel, Variabilitäts-koeffizient	Länge, Gewicht, Zeitintervalle, Testrohwerte bei probabilistischen modellkonformen Tests, Reaktionshäufigkeit

3.4. Skalenniveau und Statistik

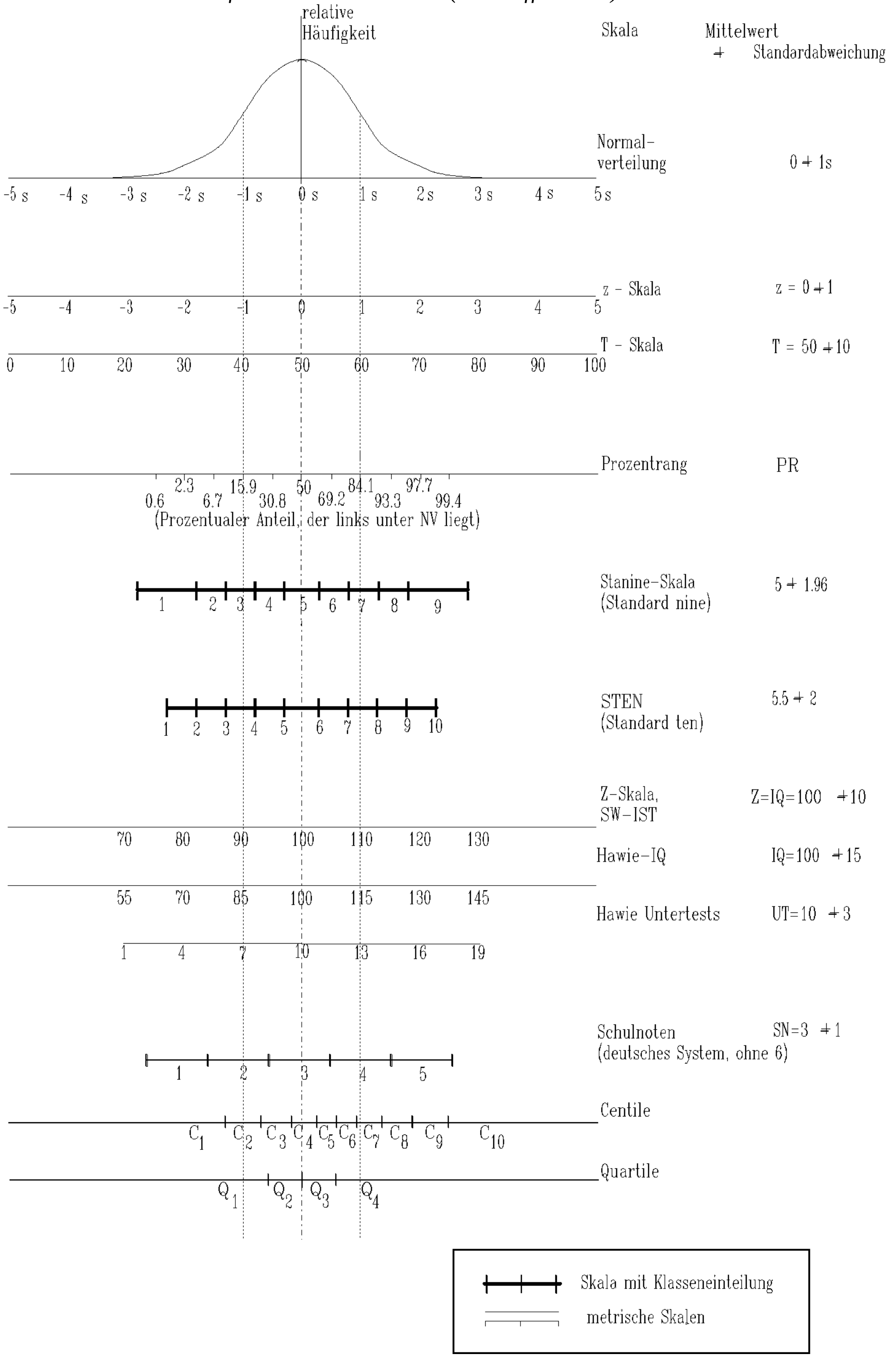
Mit statistischen Methoden kann man Zahlen analysieren. Wenn wir etwas gemessen haben, müssen wir entscheiden, auf welchem Skalenniveau sich die erhaltenen Zahlen befinden. Daraus folgt dann, welche Transformationen und auch welche statistische Methoden erlaubt sind. Oft kann man aber nicht mit Sicherheit sagen, auf welchem Messniveau sich psychologische Daten befinden.
Folgendes modifiziertes Beispiel aus Diehl und Kohr (1989; ab S.22) soll dies verdeutlichen:
Man registriert die Deutschnoten von 5 Schülern aus verschiedenen Städten: Karl hat eine 3, Gernot eine 2, Fritz eine 4, Hermann eine 1 und Friedhelm eine 5. Forscher A sagt, dass die Schüler aus verschiedenen Städten kommen, und damit nicht vergleichbar sind. Für ihn war die Messung damit auf Nominalniveau. Für Forscher A haben statistische Methoden keinen empirischen Sinn. Forscher B hingegen behauptet, dass man zumindest eine Rangreihe der Deutschleistungen bilden kann, mit Friedhelm als Bestem und Hermann als Schlechtestem (Schweizer Notensystem). Er kann jetzt z.B. den Median bestimmten (der Punkt oberhalb dessen 50% der besseren und unterhalb dessen 50%

der schlechteren Noten liegen). Er kann damit sagen, dass Fritz zu den 50% besseren Deutschschülern gehört. Forscher C geht noch weiter indem er sagt, dass auch die Abstände zwischen den Noten empirischen Sinn haben. Auf dem postulierten Intervallniveau kann nun der Mittelwert der Schulnoten berechnet werden. Untersucht man eine zusätzliche Gruppe von 5 Schülern, so kann man signifikante Unterschiede in den durchschnittlichen Schulnoten der beiden Gruppen feststellen.

3.5. Gebräuchliche Skalen für Tests

Die Rohwerte, die man als Ergebnisse der Tests erhält sind oft nur auf Ordinalskalenniveau angeordnet. Mit einem kleinen Trick "täuscht" man das Intervallskalenniveau vor, indem man die Verteilung der Rohwerte in eine Normalverteilung transformiert (Flächentransformation nach McCall, vgl. Diehl & Kohr, 1989, S. 143-148). Im Anhang der Tests finden man schliesslich eine Tabelle in der man für ermittelte Rohwerte die Normwerte nachschauen kann. Oftmals ist in diesen Tabellen noch nach diversen Subpopulationen unterschieden (z.B. Männer vs. Frauen; verschiedene Altersgruppen). Leider verwenden nicht alle Tests dieselbe Skala, wie z.B.: die Standardnormalverteilung mit Mittelwert 0 und Standardabweichung 1, sondern es gibt viele verschiedene Normskalen.
Die folgende Aufstellung vergleicht die gebräuchlichsten Skalen miteinander:

Tabelle 4: Skalen-Transformation nach Kokott (unveröffentlicht)



3.6. Uebungsaufgaben zu den Grundlagen der Messung

1. Zeichne die Mengen A und B, wobei die Menge A Personen und die Menge B Zahlen enthalten soll. Zeige eine isomorphe und einen homomorphe Abbildung der beiden Mengen ineinander.

2. Wie definieren Kranz et al. (1971) die Messung?

3. Gib eine Definition für Skalierung.