

ÜBER DEN AFFEKTWERT VON WORTDARSTELLUNGEN  
- EINE UNTERSUCHUNG AN KINDERN

von

Maria D. Simon

Juli 1966

Research Memorandum No. 4

## ÜBER DEN AFFEKTWERT VON WORTDARSTELLUNGEN - EINE UNTERSUCHUNG AN KINDERN

Die vorliegende Arbeit stellt einen Versuch dar, die Affektbedeutung von Begriffen bei jüngeren Kindern zu erfassen. Es sei vorausgeschickt, daß die Affektbedeutung eines Begriffs nur einen Teilaspekt des totalen "Bedeutungsaggregats", nämlich den gefühlsbetont-wertenden, betrifft. Schon in dem grundlegenden Werk über Semantik "The Meaning of Meaning" unterschieden Ogden und Richards (1923) zwischen der affektiven (konnotativen) und der symbolischen (denotativen) Bedeutung der Begriffe. Floyd Allport (1955) weist darauf hin, daß nur die konnotative Bedeutung meßbar ist, da sie durch skalierbare Attribute charakterisiert wird, während die denotative Bedeutung dichotom, und folglich nicht meßbar ist (d.h. ein bestimmtes "Bedeutungsaggregat" kann beispielsweise nur entweder einen Hammer oder keinen Hammer symbolisieren, es kann aber keinesfalls etwas dazwischen Liegendes kennzeichnen).

Während in den älteren Werken die Ausdrücke "affektiv" und "konnotativ" gewöhnlich wie Synonyme verwendet wurden, wissen wir heute auf Grund von faktorenanalytischen Untersuchungen mittels der Methode des semantischen Differentials von Osgood und Mitarbeitern (1957), daß in die konnotative Bedeutung sowohl vorwiegend affektiv bestimmte Attribute als auch nichtaffektive Attribute eingehen. Beim semantischen Differential werden der Versuchsperson Stimulusworte (meist Substantiva, wie etwa das Wort "Mutter") dargeboten, welche entlang einer größeren Anzahl von bipolaren Adjektivskalen mit den Noten 1 bis 7 zu bewerten sind. Aus den Interkorrelationen zwischen den Skalenwerten bestimmt man

die Dimensionen des "Bedeutungsraumes", sowie die Lokalisierung der einzelnen Begriffe in diesem Raum. Aus zahlreichen faktorenanalytischen Untersuchungen, welche Osgood durchführte (1957, 1962), geht hervor, daß der bei weitem wichtigste Faktor in der konnotativen Begriffsbestimmung der evaluative Faktor ist, welcher durch Skalen wie "schön - häßlich", "gut - schlecht", "angenehm - unangenehm" charakterisiert wird. Dieser Faktor ist für 70 % der gemeinsamen Varianz der Variablen verantwortlich. Man könnte den evaluativen Faktor auch als den "Affektfaktor" der Begriffsbestimmung bezeichnen, während die beiden andern maßgeblichen Faktoren "Potenz" (charakterisiert durch Skalen wie "groß - klein", "leicht - schwer" u. dgl.) und "Aktivität" ("schnell - langsam", "aktiv - passiv" etc.) eher die nicht-emotiven Begriffsdimensionen kennzeichnen. (Siehe Schema, Abb. 1)

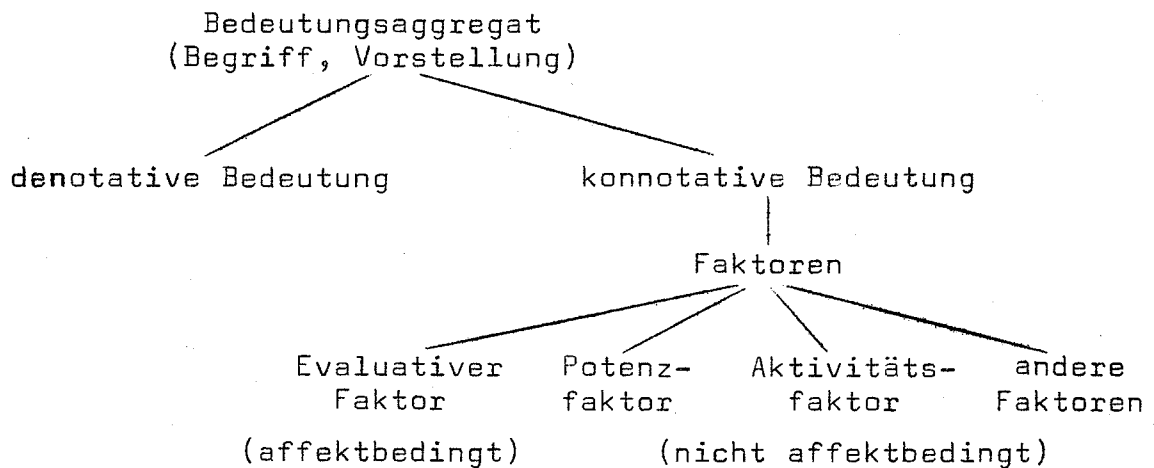


Abb. 1 Schema der Begriffsbedeutung

Untersuchungen mittels der Methode des semantischen Differentials wurden bisher fast immer nur an erwachsenen Versuchspersonen durchgeführt, und in einem Fall, wo eine Untersuchung an Schulkindern stattfand (Lilly, 1965) waren die Versuchspersonen über 8 Jahre alt und des Lesens kundig.

Der Grund für das Fehlen von Untersuchungen an Kleinkindern ist zweifellos der, daß die Methode bereits ein Maß an Abstraktionsfähigkeit erfordert, wie es bei jüngeren Kindern reifemäßig nicht zu finden ist.

Hier soll nunmehr über einen Versuch berichtet werden, die Affektbedeutung von Begriffen mittels einer Art kindertümlicher Version des semantischen Differentials zu messen, wobei in Anbetracht des begrenzten Verständnisses der Versuchspersonen die Messung nur entlang einer evaluativen Bedeutungsdimension vor sich geht. Es war unser Bestreben festzustellen,

1.) ob es überhaupt für jüngere Kinder möglich ist, mittels unserer Methode die affektive Bedeutung von Begriffen aus ihrem Lebensbereich zum Ausdruck zu bringen, bzw. was die Minimalanforderungen an Alter oder Reifestand sind.

2.) ob sich verschiedene Subgruppen (nach Alter, Geschlecht, sozialer Herkunft etc.) in ihrer Bewertung der Begriffe unterscheiden, und schließlich

3.) ob sich die Bewertung in verschiedenen Lebenssituationen als stabil erweist oder ob situationsbedingte Wertverschiebungen eintreten.

## METHODE

### Versuchspersonen

Die Versuchspersonen standen im Alter zwischen 3 und 9 Jahren und wurden in 2 verschiedenen Situationsgruppen erfaßt. Diese waren A.) eine Tagesheimgruppe, bestehend aus 102 gesunden Kindern aus Wiener städtischen Kindergärten und Schülerhorten, und B.) eine Spitalsgruppe von 50 Kindern, die sich zum Zweck einer Operation im Krankenhaus befanden.

### Testmaterial

Das Material unterschied sich von dem für Erwachsene verwendeten schon durch die Art der Begriffsrepräsentation, die nicht rein verbal war, sondern durch Bilddarstellung unterstützt wurde. Der von uns für diesen Zweck entworfene Bildwertungstest besteht aus 70 farbigen Bildkärtchen - ähnlich wie Bilderlottokarten - auf denen Begriffe aus der kindlichen Erlebniswelt dargestellt sind. Die Bilder sind von der Versuchsperson in 5 Wertränge einzuordnen.

Die Auswahl der Begriffe war ursprünglich einigermaßen willkürlich. Es wurden solche gewählt, von denen aus der pädagogischen und klinischen Erfahrung bekannt ist, daß sie häufig positive, negative oder ambivalente Affektreaktionen bei Kindern hervorrufen. Krankheits- und Spitalsbegriffe waren in der Liste deshalb stark vertreten, weil der Plan bestand, die Wirkung eines Spitalsaufenthaltes auf den Affektwert von "Spitalsbegriffen" zu untersuchen. Nach einigen Vorversuchen an Kindergartenkindern wurden aus der ursprünglichen Liste Wörter ausgeschieden, die nicht streuten, bzw. mit beibehaltenen Wörtern sehr hoch korrelierten. Ebenso wurden Wörter ausgeschieden, die von vielen Kindern nicht verstanden wurden; gewisse "Spitalsbegriffe" (wie Narkose oder Operation), die auch in letztere Gruppe fielen, wurden jedoch beibehalten, da wir annahmen, daß diese Worte wohl für viele gesunde Kinder bedeutungslos sind, jedoch bei Spitalerfahrung Bedeutung annehmen.

Die in der endgültigen Form des Tests verwendeten 70 Bildbegriffe waren:

|                 |            |         |
|-----------------|------------|---------|
| Maus            | Krampus    | beten   |
| Rauchfangkehrer | Kinderheim | Spital  |
| Haare schneiden | Messer     | ich     |
| krank           | Klosett    | Narkose |
| tanzen          | Bruder     | Spinne  |

|                  |              |               |
|------------------|--------------|---------------|
| Weihnachtsmann   | Kirche       | Zauberer      |
| rutschen         | Affe         | Bienen        |
| essen            | Wachmann     | Pillen        |
| Operation        | Injektion    | allein        |
| Geld             | Spinat       | Friedhof      |
| Neger            | Baby         | Hexe          |
| Apotheke         | schießen     | Teddybär      |
| Kindergarten     | Lutscher     | Tiger         |
| finster          | Hund         | Milch trinken |
| Schwester        | kaputt       | Feuer         |
| Flugzeug         | Gewitter     | angebunden    |
| Arzt             | klettern     | Mutter        |
| viele Leute      | schlafen     | Raupe         |
| Unfall           | Zahnarzt     | Fallschirm    |
| Wohnung          | Schlange     | Indianer      |
| Krankenschwester | Vater        | eingesperrt   |
| Schere           | waschen      | schwimmen     |
| Thermometer      | Rettungsauto |               |
| Krokodil         | boxen        |               |

### Testvorgang

Die praktische Durchführung des Tests ging derart vor sich, daß dem Kind vorerst 3 Schachteln geboten wurden, und zwar eine "schöne" Schachtel aus geblumtem Karton, eine "schiache" <sup>1)</sup> Schachtel aus schwarzem Karton und eine "mittlere" Schachtel aus grau-weißem Material. Der Versuchsleiter sprach das auf dem Bild dargestellte Reizwort vor der Darbietung laut aus, reichte dem Kind dann das Bild und forderte es auf, "die schönen Sachen in die schöne Schachtel, die schiachen Sachen in die schiache Schachtel, und die Sachen, die dir wurst <sup>2)</sup> sind, in die mittlere Schachtel" zu legen.

---

1) "Schiach" bedeutet im Wiener Idiom "häßlich" oder "schlecht".

2) gleichgültig

Die Anweisungen wurden mehrere Male wiederholt und kindertümlich erläutert und das Placieren geübt, bis das Kind die Instruktionen verstanden hatte. Nachdem die 70 Karten auf die 3 Schachteln aufgeteilt waren, wurden die positive und die negative Schachtel noch einmal aufgespalten, und zwar jeweils in eine "sehr schöne" und eine "bissi schöne", bzw. eine "sehr schiache" und eine "bissi schiache" Schachtel, sodaß schließlich 5 Wertrangstufen erreicht wurden. Das Ergebnis wurde jeweils auf einem Formblatt festgehalten.

Aus den Testresultaten lassen sich die individuellen Beliebtheitswerte der einzelnen Begriffe erkennen, aus welchen individuelle Indices kombiniert und Gruppenmaßzahlen errechnet werden konnten.

#### Validierung

Dem Hauptversuch gingen Vorversuche zur Überprüfung der Brauchbarkeit des Testmaterials und dessen möglichen Anwendungsbereichs voraus.

Die von uns gewählte Art der Begriffsdarstellung durch bunte Bilder eröffnete zweifellos eine Fehlerquelle. Zur besseren Validierung des Materials wäre es wünschenswert gewesen, Versuche mit verschiedenen variierten Bilddarstellungen durchzuführen. Dies war jedoch aus praktischen Gründen undurchführbar. Wir gaben den Test jedoch einer Gruppe von 30 achtjährigen Anstaltskindern (15 Buben und 15 Mädchen), sowohl in Bild- als auch in einer verbalen Version. Die Korrelationen zwischen Bild- und Verbalform verteilten sich bei den 70 Testitems um einen Mittelwert von  $r = 0.41 \pm 0.18$ .

Die Zuverlässigkeit wurde an 50 Kindern einer Kindergartengruppe im Alter von 3 bis 6 Jahren überprüft. Diese Kinder wurden jeweils 2 mal mit einem Zeitintervall von 2 Wochen getestet. Die Korrelationen variierten bei den 70 Items von 0,5 bis 0,81 bei einem Mittelwert von  $r = 0.36 \pm 0.19$ . Koeffizienten im gleichen Bereich fanden Lilly (1965), Kahne-

mann (1963) und Norman(1959) bei älteren Kindern und Hochschülern.

Da es klar ist, daß nicht beliebig junge Kinder mit dem Testmaterial im Sinne der Aufgabenstellung umgehen können, galt eine weitere Voruntersuchung der Feststellung der unteren Altersgrenze, bzw. des Reifestandes, der zu einer sinngemäßen Bewältigung der Testaufgabe erforderlich ist. (Für eine detaillierte Darstellung dieses Teils der Untersuchung, siehe Simon & Formanek, 1964). Von der Annahme ausgehend, daß die Fähigkeit, Bildsymbole zu werten, möglicherweise funktional mit der Leistungsfähigkeit im begrifflichen Denken in Beziehung steht, ermittelten wir den Zusammenhang zwischen den Leistungen im Bildwertungstest und einem Begriffsbildungstest bei einer Versuchsgruppe von 48 Kindern zwischen 3 und 9 Jahren. Die Leistungen im Bilderwerten wurde durch einen Beständigkeitskoeffizienten gemessen (dieser wurde analog einer Produkt-Momentkorrelation aus den Bewertungen der Bilder bei zwei Testungen errechnet), die der Begriffsbildung aus einem Punktskore; die Zahl der erreichbaren Punkte lag, je nach Abstraktionsleistung zwischen 0 und 5. Tabelle 1 zeigt den Zusammenhang zwischen den Leistungen.

Tabelle 1. Zusammenhang zwischen Güte der Abstraktionsleistung (Begriffsbildungsskore) und Beständigkeit im Werten (Bildwertungstest)

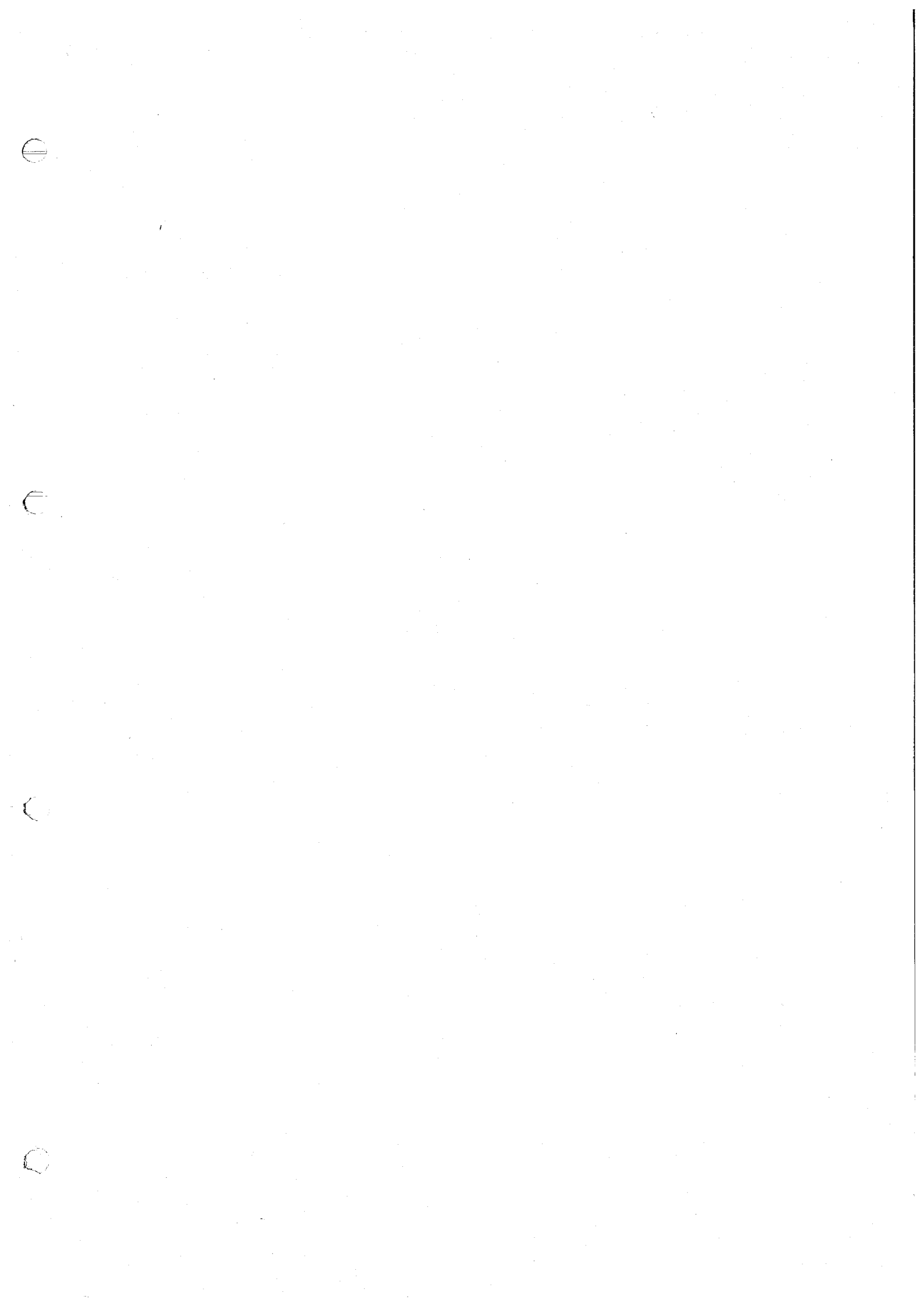
| N   | 10  | 18  | 8   | 8   | 4   |
|---|-----|-----|-----|-----|-----|
| Begriffsbildungsskore (in Punkten)                            | 0   | 1   | 2   | 3   | 4-5 |
| Mittelwert des Beständigkeitskoeffizienten (Bildwertungstest) | .25 | .49 | .55 | .59 | .60 |

Wir sehen den stärksten Anstieg in der Zuverlässigkeit des Wertens zwischen den Kindern, welche einen Begriffsbildungsskore von 0 und jenen, die einen Skore von 1 erreichen. Danach flacht die Leistungskurve im Wertens ab, und wir können feststellen, daß bei einem Begriffsbildungsskore von 2 und darüber der Leistung im begrifflichen Denken kein wesentlicher Zuwachs mehr in der Leistung des Wertens entspricht. Demnach hat der Bildwertungstest bei Kindern, die einen Begriffsbildungsskore von weniger als 1 erzielen, nur geringe Reliabilität. Wir konnten feststellen, daß im allgemeinen erst Kinder von etwa 4 1/2 Jahren aufwärts einen Begriffsbildungsskore von 1 oder darüber erreichen und schlossen daraus, daß der Bildwertungstest bei noch jüngeren Kindern keine brauchbaren Ergebnisse liefert. Weiters konnten wir feststellen, daß innerhalb der von uns untersuchten Altersgruppe der Fünf- bis Neunjährigen der Begriffsbildungstest einen besseren Maßstab für die Beständigkeit des Wertens liefert als das chronologische Alter der Versuchsperson. Dies ist aus den partiellen Korrelationen zwischen Alter, Begriffsbildungsskore und Beständigkeitskoeffizient im Wertungstest erkennbar. Die Korrelation zwischen Beständigkeit im Wertens und Lebensalter, unter Konstanthaltung des Begriffsbildungsskores, beträgt 0.09 (n.s.), zwischen Begriffsbildungsskore und Alter, bei Konstanthaltung des Beständigkeitskoeffizienten, 0.11 (n.s.), die Korrelation zwischen Begriffsbildungsskore und Beständigkeitskoeffizienten, bei Konstanthalten des Alters, jedoch 0.46 ( $p < 1\%$ ).

## ERGEBNISSE

### Wertrangordnung

Aus den Testresultaten wurde für jeden Begriff ein Beliebtheitswert errechnet. Dies ist der Mittelwert der Einstufungen von 1 (sehr schön) bis 5 (sehr schiach). Die Einstufung 3 entspricht hier dem Nullpunkt. Abb. 2 zeigt die



Beliebtheitswert-Rangordnung für die gesamte Tagesheimgruppe von 102 Kindern (Medianwerte und Interquartilspannen).

Ranglisten für zwei verschiedene Tagesheime korrelierten 0.87, waren einander also weitgehend ähnlich. In der vorliegenden Rangliste finden wir zwei Drittel der Begriffe positiv (unter Medianwert 3) und ein Drittel negativ eingestuft.

Aus der Rangliste ist ersichtlich, daß der beliebteste Begriff "Weihnachtsmann" (mit einem mittleren Beliebtheitswert von 1.1) und der unbeliebteste "Hexe" (4.65) ist. Es läßt sich erkennen, daß die Abstände zwischen den Mittelwerten aneinandergrenzender Begriffe ungleich groß sind; während das "Mittelfeld" eng besetzt ist, liegt die Bewertung der Extreme weiter auseinander. Auch die Streubreite der einzelnen Begriffe ist aufschlußreich. So findet man, daß beispielsweise bei den "Familienbegriffen", "Baby" durch starken Konsensus zu einem "guten" Wort gestempelt wurde, während die Meinungen über "Mutter", "Schwester", "Vater" und "Bruder" weit mehr auseinandergehen. Hätte man angenommen, daß religiöse Begriffe durch affektive Querverbindungen mit Vorstellungen über Sünde, Hölle und Tod etc. auch eine größere Anzahl negativer Reaktionen hervorrufen würden, so wäre dies bei der vorliegenden Stichprobe nicht bestätigt worden. Religiöse Symbole wurden dermaßen einhellig positiv bewertet, daß nach den Vortestungen die Begriffe "Engel" und "Christbaum" eliminiert wurden, da sie praktisch mit "Weihnachtsmann", "Kirche" und "beten" kovariierten. Selbst der "Friedhof" scheint relativ hoch bewertet, vermutlich weil die Grabbesuche zu Allerheiligen bei vielen Kindern eher eine assoziative Verbindung zu einem vergnüglichen Fest als zu der Idee des Sterbens hervorrufen. (Ein Kind antwortete auf die Frage, ob es wisse, was "Friedhof" bedeute: "Das ist, wo die schönen Gräberln sind").

"Milch" und "Spinat" wurden, entgegen unseren Erwartungen, von den Kindern höher bewertet als "essen" im allgemeinen.

Begriffe, die den Kindern der Tagesheimgruppe nicht geläufig waren, wie z.B. "Operation", "Pillen" und "Narkose"

befinden sich im mittleren Teil der Skala und weisen nur geringe Streuung auf.

Bei den negativ bewerteten Begriffen fallen die vielen Stereotype auf. So hat beispielsweise der "Neger" einen bedeutend tieferen Rang als der "Indianer", obwohl Wiener Kinder kaum jemals einen Neger zu Gesicht bekommen haben, ebensowenig wie den auf der Rangliste nächst dem Neger placierten "Rauchfangkehrer". Auch werden gewisse Tiere - in erster Linie kriechende, wie Raupe und Schlange - von der Mehrzahl der Kinder abgelehnt. Bei anderen Tieren, wie "Krokodil" und "Tiger", finden wir eine große Variationsbreite, also weit geteilte Meinungen. Die verbreitete Ansicht, daß Mäuse Ablehnung hervorrufen, wird von unserer Versuchsgruppe nicht geteilt.

#### Die Bewertung durch verschiedene Subgruppen

Die Bewertung der einzelnen Begriffe wurde auch für verschiedene Untergruppen innerhalb der Tagesheimgruppe berechnet, und in allen Fällen ergab sich eine annähernd symmetrische Verteilung der Bewertungen um einen Mittelwert von etwa 2.7. Das trifft sowohl für zwei verschiedene Tagesheimgruppen als auch für drei verschiedene Altersgruppen zwischen 3.9 und 8.9 zu. Die Mittelwertunterschiede waren in allen Fällen nicht signifikant. Auf die hohe Korrelation zwischen den Wertranglisten der beiden Tagesheime ist bereits hingewiesen worden.

Des weiteren wurden Gruppenunterschiede in der Bewertung der einzelnen Begriffe mittels des Chi-Quadrattests untersucht, und zwar nach folgenden Gesichtspunkten:

- a) soziale Schicht <sup>1)</sup> (Unterschicht vs. Mittel- und Oberschicht),

---

1) Die soziale Klassifizierung erfolgte nach den Kriterien der Sozialwissenschaftlichen Forschungsstelle der Universität Wien (Jugend in Wirtschaft u. Gesellschaft, 1962) in folgender Weise: die Oberschicht umfaßt Versuchspersonen, deren Eltern höhere Beamte, höhere Angestellte oder Unternehmer sind oder höhere Intelligenzberufe ausüben, die Mittelschicht entstammt den Kreisen der mittleren Beamten und Angestellten, dem Gewerbe, dem Kleinhandel oder der niedrigeren Intelligenzberufe; der Unterschicht gehören Kinder niedriger Angestellter und Arbeiter an.

- b) Alter (Altersgruppe A vs. C),
- c) Geschlecht, und
- d) verhaltensgestört vs. nicht verhaltensgestört <sup>1)</sup>.

Tab. 2, 3, 4 und 5 zeigen, welche Begriffe eine signifikant unterschiedliche Bewertung durch verschiedene Subgruppen erfahren haben.

Tabelle 2. Signifikant unterschiedlich bewertete Begriffe nach sozialer Schicht (untere soziale Schicht N = 71; Mittel- und Oberschicht N = 31)

| Begriff         | Chi Quadrat | p kleiner als | Richtung d. Unterschieds  |
|-----------------|-------------|---------------|---------------------------|
| Spinne          | 9.66        | 5 %           | keine klare Richtung      |
| Haare schneiden | 11.29       | 2 %           | beliebter b. Unterschicht |
| tanzen          | 11.19       | 2 %           | beliebter b. Unterschicht |
| allein          | 10.68       | 5 %           | keine klare Richtung      |
| Klosett         | 12.23       | 2 %           | beliebter b. Unterschicht |
| Lutscher        | 10.07       | 5 %           | keine klare Richtung      |

df = 4

Tabelle 3. Signifikant unterschiedlich bewertete Begriffe nach Altersklassen (Altersklasse A: 3.9 bis 5.5, Altersklasse C: 6.6 bis 8.9)

| Begriff     | Chi Quadrat | p kleiner als |
|-------------|-------------|---------------|
| Thermometer | 9.03        | 5 %           |
| Rettung     | 10.16       | 5 %           |
| eingesperrt | 9.66        | 5 %           |

df = 4. Richtung des Unterschiedes: In allen Fällen war der Begriff bei den älteren Vpn weniger beliebt. Altersklasse A: N=41  
Altersklasse B: N=22

<sup>1)</sup> Die Klassifizierung verhaltensgestört - nicht verhaltensgestört wurde auf Grund von Beurteilungen der Gruppenleiterinnen und Direktorinnen der Tagesheime erstellt.

Tabelle 4. Signifikant unterschiedlich bewertete Begriffe nach Verhalten (verhaltensgestört N = 26, nicht verhaltensgestört N = 76)

| Begriff   | Chi Quadrat | p kleiner als | Richtung des Unterschiedes                            |
|-----------|-------------|---------------|---|
| Operation | 9.90        | 5 %           | Beliebter bei verhaltensgestörten Vpn.                |
| essen     | 11.98       | 5 %           | mehr extreme Bewertungen bei verhaltensgestörten Vpn. |
| Gewitter  | 12.1        | 2 %           | mehr extreme Bewertungen bei verhaltensgestörten Vpn. |

df = 4

Tabelle 5. Signifikant unterschiedlich bewertete Begriffe nach Geschlecht (Knaben N = 55, Mädchen N = 47)

| Begriff  | Chi Quadrat | p kleiner als | Richtung des Unterschiedes         |
|----------|-------------|---------------|------------------------------------|
| Mutter   | 10.13       | 5 %           | beliebter bei Mädchen              |
| Polizist | 10.04       | 5 %           | beliebter bei Buben                |
| boxen    | 9.18        | 5 %           | beliebter bei Buben                |
| klettern | 10.64       | 5 %           | beliebter bei Buben                |
| Rettung  | 10.45       | 5 %           | mehr extreme Bewertungen bei Buben |
| Krokodil | 13.08       | 2 %           | beliebter bei Buben                |
| Lutscher | 9.66        | 5 %           | beliebter bei Mädchen              |
| Indianer | 16.14       | 1 %           | beliebter bei Buben                |
| tanzen   | 13.7        | 1 %           | beliebter bei Mädchen              |
| Baby     | 16.19       | 1 %           | beliebter bei Mädchen              |
| Tiger    | 15.38       | 1 %           | beliebter bei Buben                |
| Schlange | 14.44       | 1 %           | beliebter bei Buben.               |

df = 4

Bei der Beurteilung der Bewertungsunterschiede muß in Betracht gezogen werden, daß für jeden Gruppenvergleich 70 Chi-Quadratanalysen durchgeführt wurden, wobei zu erwarten steht,

daß rein zufällig etwa drei bis vier Werte bei 5 % signifikant aufscheinen. Man darf daher annehmen, daß die unterschiedlichen Bewertungen zwischen den Angehörigen verschiedener sozialer Klassen (betreffend 6 Begriffe), Altersgruppen (3 Begriffe) oder Verhaltensgruppen (3 Begriffe) unter Umständen auf Stichprobenfehlern beruhen, umso mehr als die Bedeutung der Unterschiede nicht immer interpretierbar ist (wie beispielsweise, wenn die Angehörigen einer Gruppe bei gewissen Begriffen die extremen Wertpositionen vorziehen, während die Vergleichsgruppe mittlere Bewertungen vorzieht). Die meisten und eindeutigsten Gruppenunterschiede (12 Begriffe) finden wir zwischen den beiden Geschlechtern: man darf auf Grund der Ergebnisse sagen, daß es anscheinend bereits bei Vorschulkindern ein "männliches" und ein "weibliches" Weltbild gibt.

#### Unterschiede zwischen den Situationsgruppen

Wie einleitend erwähnt wurde, war es beabsichtigt, den Einfluß von Situationsfaktoren auf die Bewertung der Begriffe zu untersuchen. Es standen 2 Versuchsgruppen für den Situationsvergleich zur Verfügung, nämlich die gesunde Tagesheimgruppe mit 102 Kindern und die Gruppe der "Spitalskinder" mit 50 Kindern. Von den "Tagesheimkindern" waren 50 Kinder zweimal mit dem Bildwertungstest untersucht worden; von der Spitalsgruppe wurden alle 50 Kinder sowohl vor als auch nach einer Operation getestet, und von diesen Kindern wurden 34 noch ein drittes Mal anläßlich eines Hausbesuchs etwa 3 Monate nach der Rückkehr aus dem Krankenhaus getestet.<sup>1)</sup>

Zur leichteren Durchführung von individuellen und Gruppenvergleichen suchten wir die 70 separaten Beurteilungen, aus denen jedes individuelle Protokoll besteht, zu einem Index zu kombinieren, sodaß jede Versuchsperson durch nur eine Maßzahl -

<sup>1)</sup> Eine ausführliche Darstellung der Spitalsuntersuchung, unter Berücksichtigung von Aspekten, welche nicht direkt mit dem Bildwertungstest zusammenhängen, erscheint separat (Simon u. Strotzka, im Erscheinen).

statt deren 70 - charakterisiert werden konnte. Dieser Bewertungsindex wird errechnet, indem alle Begriffe aus einer der beiden positiven Kategorien mit je einem Pluspunkt, die Begriffe aus den beiden negativen Kategorien mit je einem Minuspunkt, und die Begriffe der mittleren Kategorie überhaupt nicht bewertet werden. Mit andern Worten, der Index drückt etwas über die Tendenz der Versuchsperson aus, positiv, negativ oder neutral (ohne Rücksicht auf begriffliche Inhalte) zu urteilen.

Die Indexzahlen der gesamte Tagesheimgruppe (N = 102) rangierten zwischen den Extremen von minus 17 bis plus 48 bei einem Mittelwert von  $17.4 \pm 14.6$ . Tabelle 6 zeigt die Mittelwerte und Standardabweichungen der Bewertungsindices von 50 zweimal getesteten Tagesheimkindern (diese Kinder sind ein Teil der Gesamttagesheimgruppe) und von ebenfalls zweimal getesteten Spitalskindern.

Tabelle 6. Vergleich der Bewertungsindices der Tagesheimgruppe und der Spitalsgruppe mittels t-Test

|  | Tagesheimgruppe<br>Alter 3.9-8.9<br>N=50 |                | Spitalsgruppe<br>Alter 4.0-8.1<br>N=50 |                |
|--|--|----------------|--|----------------|
|  | Test 1<br>(I)                            | Test 2<br>(II) | Test 1<br>(III)                        | Test 2<br>(IV) |
| Mittelwert<br>u. Standard-<br>abweichung | 18.16±14.2                               | 15.07±14.8     | 7.9±17.1                               | 4.7±14.6       |

Signifikanz der Mittelwertunterschiede:

$$t_{I-II} = 1.23 \text{ (n.s.)}^a)$$

$$t_{II-IV} = 1.65 \text{ (n.s.)}^a)$$

$$t_{I-III} = 6.57 \text{ (p < 1\%)}^b)$$

$$t_{I-IV} = 9.35 \text{ (p < 1\%)}^b)$$

$$t_{II-III} = 4.44 \text{ (p < 1\%)}^b)$$

$$t_{II-IV} = 7.05 \text{ (p < 1\%)}^b)$$

a) t-Test für korrelierende Stichproben, df = 49

b) t-Test für unabhängige Stichproben, df = 98



[Faint, illegible text or smudge]

Man sieht, daß wohl innerhalb der beiden großen Gruppen keine signifikanten Unterschiede, zwischen ihnen jedoch stark signifikante Unterschiede bestehen: die Bewertungsindices der Spitalskinder sind bedeutend niedriger als die der Kinder ohne Spitalerfahrung.

Es war nun von Interesse festzustellen, daß die Bewertungsindices, welche anlässlich der dritten Testung bei einem Hausbesuch einige Monate nach der Spitalsentlassung gefunden wurden, keineswegs zum Niveau der Gruppe ohne Spitalerfahrung zurück tendierten, sondern nur geringfügig über dem Wert der postoperativen Testung lagen, nämlich bei einem Mittelwert von  $6.9 \pm 10.5$ . Es lag die Vermutung nahe, daß die Ursache der gedrückten Bewertungsindices in einer andauernden Abwertung von Begriffen zu suchen sei.

Eine Prüfung der Bewertung einzelner Begriffe im Verlaufe von 3 Testungen (präoperativ, postoperativ, Hausbesuch) ergab nun, daß keineswegs alle Begriffe von einer Abwertung betroffen waren, sondern daß die Wertverschiebung durchaus selektiv verlaufen war. Begriffe, welche unter dem Sammelnamen "Angstbegriffe" (stechende und beißende Tiere, bedrohliche Objekte und Gestalten) und "Spitalsbegriffe" subsumiert werden können, waren von Testung zu Testung abgewertet, hingegen Begriffe, die mit Heim und Familie zusammenhingen, progressiv aufgewertet worden. Die zur Abwertung gelangten Begriffe waren in der Wortliste stärker vertreten als die aufgewerteten, was in den gedrückten Bewertungsindices seinen Niederschlag fand. Abb. 3 und 4 zeigen, wie die Bewertung von "Spitalsbegriffen" und "Familienbegriffen" zwischen Tagesheimgruppe und Spitalsgruppe differiert.

Man sieht, daß die Spitalskinder die "Familienbegriffe" durchwegs positiver, die Spitalsbegriffe (mit Ausnahme der "Pflegerin") negativer beurteilen als die Tagesheimkinder. Der Unterschied ist bei 5 % signifikant (Vorzeichentest).

Das Schicksal der Begriffe im Verlaufe der drei Testungen an der Spitalsgruppe (präoperativ, postoperativ und Hausbesuch) zeigt Abb. 5.



Es wurden für jede Testung ein "Familienbegriffsindex" und ein "Spitalsbegriffsindex" errechnet, und zwar für jede Versuchsperson separat durch Ermittlung des Durchschnitts der Beliebtheitswerte für die 6 Familienbegriffe (siehe Abb. 3) und der 10 Spitalsbegriffe (siehe Abb. 4). Hier erweist sich, daß die Minderbewertung der Spitalsbegriffe progressiv war und sogar weitere Fortschritte machte, nachdem die Kinder nach Hause zurückgekehrt waren. Gleichzeitig setzte sich nach der Rückkehr die progressive Aufwertung von Heim und Familie fort.

#### DISKUSSION

Schon eingangs wurde festgestellt, daß keineswegs der Anspruch gemacht wird, mittels einer eindimensionalen Skala die volle "Bedeutung" von Begriffen erfassen zu können. Andererseits konnte darauf verwiesen werden, daß selbst bei Anwendung eines multidimensionalen Meßvorgangs und bei faktorenanalytischer Auswertung sich der "evaluative" Faktor als der weitaus wirksamste in der Begriffsbestimmung erwiesen hat. Es dürfte eine vertretbare Annahme sein, daß die von uns verwendete Skala "schön - schiach" für jüngere Kinder tatsächlich die meisten "evaluativen" Attribute miteinschließt.

Die Ergebnisse der vorliegenden Untersuchung deuten darauf hin, daß Vorlieben und Abneigungen (und somit "Bedeutungen" im emotiven Sinn) sich bereits frühzeitig herausbilden und daß sie eine gewisse Allgemeinverbreitung haben. Starker Konsensus fand sich bei unserer Stichprobe zwischen verschiedenen Altersgruppen (zwischen 4 und 9 Jahren), sozialen Gruppen und Verhaltensgruppen, jedoch in viel geringerem Maß zwischen den Geschlechtern. Die Unterschiede in der Bewertung von Begriffen machen es möglich, gewisse Hypothesen zu erhärten

oder zu widerlegen. So widerspricht es beispielsweise der psychoanalytischen Lehrmeinung, wenn bei unserem Versuch sowohl von Buben als auch von Mädchen die Mutter höher als der Vater gewertet wird, und bei Mädchen noch signifikant höher als bei Knaben.

Es ergibt sich auch die Frage, ob die für unsere Stichprobe gefundenen Werte für Kinder im allgemeinen zutreffen oder ob sie kulturspezifisch sind. Dies wäre selbstverständlich nur durch interkulturelle Vergleiche schlüssig zu beantworten, aber unser Material stützt die Hypothese, daß möglicherweise gewisse Begriffe häufig mit universellen Angsterregern assoziiert werden, wie etwa dunkle Personen (Neger, Rauchfangkehrer, Krampus) und kriechende Tiere.

Weiters gab unsere Untersuchung darüber Aufschluß, daß die Begriffsbedeutung, selbst falls sie im Sinne der Universalität vorstrukturiert sein sollte, keinesfalls von Erfahrungsinhalten unabhängig ist. Wir sahen, daß die Bekanntschaft mit einem Krankenhaus eine weitgehende Wertverschiebung bei den untersuchten Kindern mit sich brachte.

Schließlich brachte die Untersuchung auch einige Aufschlüsse über die Entwicklungsstufe, in der Kommunikations-symbole ihre Affektwertigkeit zugeordnet erhalten. Ohne Zweifel empfinden Kinder zu den Objekten ihrer Umwelt Zu- oder Abneigung, lange bevor sie imstande sind, den Affekt mit seiner symbolischen Repräsentation zu verknüpfen. Wir konnten nun durch die parallele Untersuchung mit Bildwertungstest und einem Begriffsbildungstest zeigen, daß eine verlässliche Wertzuordnung zu den Begriffen erst etwa ab dem 4 1/2. Lebensjahr stattfindet.

Da sich der Bildwertungstest, selbst in seiner gegenwärtigen primitiven Form, als recht ergiebiges Forschungsinstrument für Gruppenuntersuchungen und Vergleiche erwiesen hat, bestünde noch die Möglichkeit zu explorieren, ob er auch individualdiagnostisches Potential hat. Ein "Bewertungsindex",

wie der von uns versuchte, der alle Items zu einer Maßzahl ohne Rücksicht auf Inhalt vereinigt, dürfte für Individualdiagnosen unbrauchbar sein. Denn obwohl sich bei unserem Versuch die Mittelwerte der Indices signifikant bei Tagesheimgruppe und Spitalsgruppe unterschieden, zeigte die Analyse der Bewertung der einzelnen Begriffe, daß der Gesamtindex vom Sinnzusammenhang der einzelnen Items nicht unabhängig ist: die "Spitalsbegriffe" und die "Familienbegriffe" kovariierten jeweils in entgegengesetzter Richtung, und die gedrückten Bewertungsindices der Spitalskinder waren letzten Endes nur eine Folge der starken Anreicherung der Testbatterie mit Bildbegriffen, die Spitals- oder Angstkonnotationen haben.

Andererseits wäre es sinnvoll zu überprüfen, ob einem Wertprofil diagnostische und prädiktive Möglichkeiten innewohnen. So ersieht man beispielsweise aus der in Abb.2 dargestellten Wertrangordnung, was die typischen Bewertungen und der mittlere Wertungsbereich für gewisse Begriffe durch eine Gruppe gesunder Wiener Kinder waren. Es sollte bei entsprechender Eichung möglich sein, festzustellen, ob ein Kind ein normales oder ein abweichendes Wertprofil aufweist, und unter welchen Bedingungen abweichende Profile auftreten.

## ZUSAMMENFASSUNG

Es wurde der Versuch unternommen, die Affektbedeutung von Begriffen bei jüngeren Kindern mittels eines für diesen Zweck entworfenen Bildwertungstests zu messen. Die Aufgabe der Versuchspersonen bestand darin, 70 bildlich dargestellte Begriffe aus der kindlichen Lebenssphäre jeweils einer von 5 Wertstufen von "sehr schön" bis "sehr schiach" (sehr häßlich) beizuordnen. Die Versuchspersonen standen im Alter von 3 bis 9 Jahren und wurden in 2 verschiedenen "Situationsgruppen" erfaßt. Die Gruppen waren 102 gesunde Kinder aus Tagesheimstätten (Tagesheimgruppe) und 50 hospitalisierte Kinder (Spitalsgruppe).

Auf Grund einer Voruntersuchung wurde festgestellt, daß der Test nur für Kinder über etwa 4 1/2 Jahren valid ist.

Die Testresultate der Hauptuntersuchung wurden in folgender Weise ausgewertet:

1) Es wurde für die Tagesheimgruppe eine Wertrangordnung der 70 Begriffe erstellt, und die Bewertung der einzelnen Begriffe durch verschiedene Untergruppierungen (nach Alter, sozialer Herkunft, Geschlecht etc.) miteinander verglichen. Dabei erwies sich, daß die Wertrangordnungen von 2 verschiedenen Tagesheimen weitgehend konform waren; bei den erwähnten Untergruppierungen fanden sich die meisten und signifikantesten Unterschiede in den Bewertungen zwischen den beiden Geschlechtern.

2) Es wurde weiters versucht, die 70 individuellen Skores zu einer Maßzahl - dem Bewertungsindex - ohne Rücksicht auf begriffliche Inhalte zu kombinieren. Es wurde festgestellt, daß die Bewertungsindices der Spitalsgruppe signifikant niedriger waren als die der Tagesheimgruppe.

3) Die Unterschiede zwischen Tagesheimgruppe und Spitalsgruppe wurden hinsichtlich der Bewertung der einzelnen Items einer weiteren Analyse unterzogen; ferner wurde das Schicksal von sinnverwandten Begriffen ("Familienbegriffen", "Spitalsbegriffen") im Verlaufe von 3 Testungen an der Spitalsgruppe

untersucht. Es zeigte sich, daß die Spitalsgruppe als Ganzes "Spitalsbegriffe" niedriger, "Familienbegriffe" höher als die Tagesheimgruppe beurteilte; und daß bei der Spitalsgruppe im Verlaufe von drei Testungen, welche jeweils vor der Operation, nach der Operation und bei einem späteren Hausbesuch stattfanden, die "Spitalsbegriffe" progressiv abwertet, die "Familienbegriffe" progressiv aufgewertet wurden.

Es wird festgestellt, daß man die Methode des Bildwertungstests als eine kindertümliche Version des semantischen Differentials ansprechen könne, mit deren Hilfe es möglich ist, zu definieren, welche "Affektbedeutung" einzelnen Begriffen zukommt. Die Methode hat gezeigt, daß innerhalb homogener Kindergruppen ein weitgehend konformes "Weltbild" herrscht, das allerdings bis zu einem gewissen Grad zwischen Buben und Mädchen variiert. Weiters konnte gezeigt werden, daß situative Faktoren - wie Spitalserfahrung - Verschiebungen in diesem Weltbild mit sich bringen, die noch lange nach dem spezifischen Erlebnis (in unserem Falle noch zur Zeit eines Hausbesuchs mehrere Monate nach der Rückkehr aus dem Krankenhaus) nachwirken.

Es wird erwogen, ob der Methode Potential für ein individualdiagnostisches Instrument innewohnt. Es sollte nach entsprechender Eichung eines Normalbereichs der Wertungen möglich sein festzustellen, ob ein Individuum ein normales oder ein abweichendes Wertprofil aufweist.

## SUMMARY

An experiment for the assessment of the affective meaning of concepts in young children by means of a Picture Evaluation Technique is presented. Subjects, aged 3 to 9, rated 70 picture concepts on a 5-point bipolar "like-dislike" scale. The subjects were drawn from two different "situation groups", one group consisting of 102 healthy children from Vienna city nursery schools and day centres (Day Care Group), the other, of 50 children in the same age range, hospitalized for minor surgery (Hospital Group). In a preliminary experiment it was found that the level of performance with the Picture Evaluation Technique correlates highly with developmental status in the area of concept formation, and that consequently the technique is only applicable to children from approximately 4 1/2 years of age.

Analysis of data from the main experiment yielded

- (1) Rank order for 70 concepts by the total Day Care Group and by various subgroupings. It was found that rankings from two different day centres, and rankings by different groups by age, social class or behavioural adjustment were substantially similar. The most numerous, consistent and significant differences in evaluation were found between the two sexes.
- (2) For each individual, scores on the 70 test items were combined into an "Evaluation Index" without regard to item content. It was found that Evaluation Indices for the Hospital Group were significantly lower than for the Day Group.
- (3) The differences in the evaluation of individual items were further analysed both in respect of the two situational groups (Day Group and Hospital Group); and with regard to the fate of conceptually related items in the course of three individual testings on members of the Hospital Group: these testings took place before an operation, after operation, and

finally during a home visit some time after release from hospital. It was found that the Hospital group as a whole rated hospital-related items significantly lower, family-related items significantly higher than the Day Group. Within the Hospital Group, hospital-related items were progressively devaluated and family-related items progressively appreciated in the course of the three testings. It is of interest that the tendency towards continuing depreciation of hospital items and appreciation of family items persisted even after release from hospital.

The Picture Evaluation Technique is considered as a simplified version of the Semantic Differential, by means of which it is feasible to explore affective meanings of concepts even in young children. The present experiment has yielded evidence of a strongly conforming "world view" within a homogenous group of young subjects, even though there appears to exist already in some areas a "feminine" and a "masculine" outlook in keeping with the subjects' sex. Further it could be shown that situational factors profoundly change concept evaluations, and that those changes appear to persist even long after the causal event.

Finally the thought is presented that the Picture Evaluation Technique might be potentially useful for the diagnosis of deviant value profiles.

## LITERATUR

- Allport, F.H., 1955. Theories of perception and the concept of structure. New York: Wiley
- Kahnemann, D., 1963. The semantic differential and the structure of inferences among attributes. Amer. J. Psychol., 76, 554-567
- Lilly, R.S., 1965. A developmental study of the semantic differential. Educational Testing Service Research Bulletin (RB-65-28) and ONR Technical Report
- Norman, W.T., 1959. Stability characteristics of the semantic differential. Amer. J. Psychol., 1959, 72, 581-584.
- Ogden, C.K. und Richards, I.A., 1923. The meaning of meaning. London: Kegan Paul.
- Osgood, C.E., Suci, G. J. & Tannenbaum, P.H., 1957. The measurement of meaning. Urbana, Ill.: U. of Ill. Press
- Osgood, C.E., 1962. Generality of affective meaning systems. Amer. Psychologist, 17, 10-28.
- Simon, M.D. & Formanek, R., 1964. Werturteilsbildung und Begriffsbildung bei jüngeren Kindern. Z. f. exper. u. angew. Psychol. XI, 623-637.
- Simon, M.D. & Strotzka, H. Eine empirische Untersuchung über die psychologische Auswirkung eines Spitalsaufenthaltes bei Kindern. Österr. Z.f.Kinderheilkunde. Im Erscheinen.