

---

---

## Originalarbeiten

---

---

Manfred Müller

### Analyse und Modifikation des Unterrichtsklimas von Berufsschulklassen

Zu Beginn der durchgeführten Klimauntersuchungen stand die Frage im Vordergrund, inwieweit das Unterrichtsklima von Berufsschulklassen verbessert werden kann, wenn es mit Hilfe eines Fragebogens gemessen und die Ergebnisse in einem Gespräch zwischen Lehrkraft und Klasse besprochen werden. Für die Überprüfung dieser Fragestellung wurden zunächst die 'Landauer Skalen zum Sozialklima, Lasso' (v. SALDERN & LITTIG, 1987) eingesetzt. Eine der ersten Erhebungen schlug jedoch fehl, weil mehrere der befragten Schüler<sup>1</sup> Items nicht beantworten konnten, die nach der Klimawahrnehmung der meisten in der Klasse fragten. Damit wurde eine intensive Analyse des in der Literatur vorfindlichen Klimakonstrukts nötig, so daß das anfängliche Ziel zunächst in den Hintergrund trat. In den Mittelpunkt rückte statt dessen zum einen eine nicht zu umgehende Neukonstruktion der 'Lasso', weil bei zukünftigen Messungen die Beantwortbarkeit der Items unbedingt sichergestellt werden sollte. Zum anderen stellte sich für die theoretische Verankerung eines neu zu konstruierenden Erhebungsinstrumentes die Frage nach einer adäquaten Klimadefinition, die in der Lage ist, auch das Gepräge der Unterrichts Atmosphäre von heterogenen Berufsschulklassen zu berücksichtigen. Dies erforderte schließlich eine unumgängliche Modifikation des weitverbreiteten Klimabegriffs von DRESMANN (1982), weil

---

<sup>1</sup> Selbstverständlich sind mit der männlichen Form, die nur aus Gründen der Vereinfachung gewählt wird, in diesem Artikel auch Schülerinnen gemeint. Dies gilt analog natürlich auch für Lehrerinnen.

sich u.a. herausstellte, daß dieser Begriff zu Itemformulierungen führte, für die eine Beantwortbarkeit nicht in jedem Fall garantiert werden kann.

Erst auf der Grundlage dieser Vorarbeiten, die auch den Schwerpunkt dieses Beitrags darstellen sollen, konnte schließlich das anfangs anvisierte Ziel in Form einer Längsschnittstudie weiterverfolgt werden. Im Rahmen dieses Artikels werden von dieser Studie lediglich die wichtigsten Hauptbefunde dargestellt, um erste empirische Hinweise auf die Validität des neukonstruierten Instruments vorzustellen. Die folgenden Ausführungen konzentrieren sich somit auf drei inhaltliche Schwerpunkte:

- Neudefinition des Klimabegriffs
- Neukonstruktion der 'Lasso'
- Längsschnittstudie mit den neu konstruierten Skalen.

#### NEUDEFINITION DES KLIMABEGRIFFS

##### Gründe für die fehlgeschlagene Klimaerhebung mit den 'Lasso'

Beim ersten Einsatz der 'Lasso', mit denen das Unterrichtsklima einer Berufsschulklasse erfaßt werden sollte, stellte sich heraus, daß mehrere Schüler Schwierigkeiten hatten, Items zu beantworten, die nicht nach der Klimawahrnehmung des jeweiligen Bearbeiters des Fragebogens, sondern danach fragten, wie *die meisten* Schüler in seiner Klasse bestimmte Klimaaspekte erleben. So wurden z.B. folgende Lasso-Items wiederholt nicht beantwortet:

- 19: Die meisten Schüler unserer Klasse hätten lieber einen anderen Lehrer.
- 87: Die meisten Schüler in unserer Klasse hätten lieber andere Klassenkameraden.
- 103: Den meisten Schülern machen die im Unterricht besprochenen Themen Spaß.

Ein anschließendes Gespräch mit der Klasse (die seit über einem Jahr in dieser Form existierte und von mir als Klassenleiter geführt wurde) zeigte, daß bei den genannten Items kein Sinn darin gesehen worden war, das zu erfassen, was „Die meisten Schüler“ denken. Ein Schüler, der Lasso-Item 87 nicht angekreuzt hatte, kommentierte seine Verweigerung so:

„Woher soll ich wissen, was *die meisten* Schüler aus meiner Klasse denken?“

Ein anderer Schüler, der den Fragebogen überhaupt nicht beantwortet hatte, schrieb auf die Fragebogen-Rückseite:

„... Ich beantworte Ihnen gerne Fragen, die sich *auf mich* beziehen. Woher soll ich aber wissen, was die meisten wollen? ...“.

Vor dem Hintergrund dieser Schwierigkeiten stellte sich die Frage, warum z.B. Lasso-Item 87 nicht so formuliert worden war, wie es die Schüler vorschlugen:

*Ich hätte lieber andere Klassenkameraden.*

Dieses Item wäre innerhalb des vorgegebenen vierstufigen Antwortmodus (1 = trifft gar nicht zu, 4 = trifft völlig zu) im Vergleich zu Item 87 eindeutig beantwortbar gewesen. Bei der Konstruktion der 'Lasso' wurde jedoch bewußt auf solche sog. „Ich-Items“ verzichtet, weil damit „... eine starke Fixierung auf den Antwortenden verbunden wäre, was dem Gruppenphänomen Sozialklima nicht Rechnung getragen hätte“ (v. SALDERN 1987, S. 222). Bei Lasso-Item 87 soll der Antwortende demzufolge nicht entscheiden, wie er persönlich diesen Einzelaspekt der Schüler-Schüler-Beziehung wahrnimmt (Ich hätte lieber ...), sondern wie ihn die meisten in der Klasse wahrnehmen („Die meisten Schüler hätten lieber ...“).

Eine tiefgehende Analyse der theoretischen Konzeption der 'Lasso' führte schließlich zu dem von DREESMANN (1982) vorgestellten Klimabegriff, der nicht nur wegen der dargestellten Beantwortungsproblematik für die hier beabsichtigten Untersuchungen ungeeignet erschien und deshalb modifiziert werden mußte, wie die weiteren Ausführungen zeigen sollen.

#### Modifikation des DREESMANNschen Klimabegriffs

In Anlehnung an die amerikanischen Sozialklimaforscher G.G. STERN (1970) und R.H. MOOS (1974) schränkte H. DREESMANN (1982, S. 45) den Klimabegriff auf diejenigen Klimawahrnehmungen ein, die *von den meisten oder sogar allen Schülern* einer Klasse geteilt werden:

Das Erleben der Lernumwelt wird erst dann als Klima bezeichnet, wenn es „...Geltung für die meisten oder sogar alle...“ Schüler hat.

Diese Begriffseinschränkung hält er deshalb für sinnvoll, weil aus seiner Sicht nur die mehrheitlich geteilten Klimawahrnehmungen auch für die Mehrheit der Schüler verhaltensrelevant sind. Aufgrund dieser Einschränkung muß DREESMANN konsequenterweise denjenigen Klassen die „Existenz“ eines Klimas ab-

sprechen, in denen die Umweltbedingungen in hohem Maße individuell erlebt werden (1982, S. 53). Ganz im Sinne dieser Einschränkung, die bei der Konstruktion der 'Lasso' übernommen wurde (v. SALDERN 1987, S. 16), ist bereits auf der Ebene der Itemkonstruktion versucht worden, nur den kollektiven Wahrnehmungsanteil zu erfassen. Die fehlgeschlagene Messung zeigt aber, daß es Berufsschüler geben kann, die über die Klimawahrnehmungen der meisten in der Klasse nichts aussagen können oder wollen, weil sie die Kognitionen, Emotionen und Motivationen der meisten u.U. nicht kennen, an der Aussagekraft solcher Daten zweifeln oder ihre eigene Wahrnehmung für bedeutsamer halten als die der *meisten* Mitschüler. Auch war nicht klar, wieviele „die meisten“ sein sollten: die absolute Mehrheit, die Zweidrittelmehrheit, 90% oder ... ?

Bei anschließenden Vorversuchen, die Klimawahrnehmung mit eigens konstruierten Skalen so zu erfassen, wie es die Schüler vorschlugen, ergab sich, daß selbst in manchen 12ten Klassen, in denen sich Schüler und Klassenleiter mehr als zwei Jahre kannten, nicht bei allen Klimamerkmale gemeinsam geteilte Wahrnehmungen meßbar waren. Ein Beispiel für fünf Subdimensionen der Schüler-Lehrer-Beziehungen soll dies verdeutlichen (Tabelle 1).

Tabelle 1

*Ergebnis einer Klimaerhebung für fünf Einzeldimensionen der Schüler-Lehrer-Beziehung*

	-	-/+	+
1. Fürsorglichkeit der Lehrkraft	5	6	10
2. Aggressionen gegen die Lehrkraft	10	1	10
3. Bevorzugung/Benachteiligung von Schülern			21
4. Autoritäre Merkmale im Unterricht der Lehrkraft	3	3	15
5. Zufriedenheit der Schüler mit der Lehrkraft	8	1	12

*Anmerkungen.* Die Abbildung zeigt das Befragungsergebnis einer 12ten Berufsschulklasse aus den Voruntersuchungen. Bei den Einzeldimensionen handelt es sich um die fünf 'Lasso'-Dimensionen der Schüler-Lehrer-Beziehungen, die hier lediglich sprachlich modifiziert wurden. „-“ bedeutet z.B. bei Dimension 1 wenig Fürsorglichkeit, bei Dimension 4 viele autoritäre Merkmale. Die Zahlen symbolisieren die Häufigkeit der Schüler, die sich für negativ, -/+ oder positiv entschieden hatten. Aus Gründen der Veranschaulichung wurden die im Bewertungsraster verwendeten zwei positiven und zwei negativen Bewertungsstufen hier jeweils zu einem einzigen Klimaurteil („-“ oder „+“) zusammengefaßt.

Geht man einmal von der Zweidrittelmehrheit als Grenzwert für die 'Existenz' eines Klimamerkmals im Sinne DREESMANN'S aus, so würden nach seiner Definition nur die Merkmale 3 und 4 'Situationsfaktoren' darstellen, die

Verhaltensrelevanz für die Schüler hätten. Die anderen wären per definitionem als Klimamerkmale nicht existent.

Damit würde offensichtlich die Bedeutung der empirisch gefundenen Lasso-Dimensionen zur inhaltlichen Beschreibung des Klimas relativiert werden, was insbesondere im Fall von Dimension 2 höchst problematisch erscheint. Denn es kann nicht davon ausgegangen werden, daß das Merkmal „Aggressionen gegen die Lehrkraft“ bei der dargestellten Ausprägung keinen Einfluß auf das Verhalten der Schülermehrheit hat. Wenn mit dem Begriff 'Klima' die typische Atmosphäre im Unterricht einer Klasse erfaßt werden soll, dann liefert das dargestellte Auswertungsergebnis z.B. für Dimension 2 vor allem bei einem beabsichtigten Datenfeedback möglicherweise wertvolle Informationen für die Klimadiagnose. Die Lehrkraft, auf die sich die obige Beispielauswertung bezog, wäre nämlich bei ihrer Beschreibung der typischen Unterrichtsatmosphäre niemals auf die Idee gekommen, Merkmal 2 als nicht-existent zu betrachten, wie sich in einem persönlichen Gespräch mit ihr herausstellte. Im Gegenteil: Vor allem dieses Einzelergebnis hatte sie als äußerst hilfreich für das Verständnis der Unterrichtsatmosphäre in dieser Klasse angesehen, die sie als belastend empfunden hatte, weswegen sie auch an einer Klimamessung im Zuge der Voruntersuchungen teilnahm.

Wenn die *Lehrkraft* diese Polarisierung der Meinungen als plausibel betrachtete und diese auch mit der soziologischen Zusammensetzung der Klasse erklären konnte, kann nicht ausgeschlossen werden, daß ihr bisheriges Verhalten von dieser Polarisierung abhängig war. Denn möglicherweise liegen vorhandene Probleme gerade in der dargestellten Polarisierung, der vor allem die Lehrkraft ausgesetzt ist, weil diese einer solchen Bedingungskonstellation (im Gegensatz zu manchen Schülern) in ihrer täglichen Arbeit nicht enttrinnen kann. Auch wenn es keinen mehrheitlich geteilten Wahrnehmungsanteil gibt (und damit nach DREESMANN kein verhaltensrelevanter Situationsfaktor vorliegt), kann eine Verhaltensrelevanz zumindest für die Lehrkraft nicht ausgeschlossen werden. Dies ist vor allem deshalb bedeutungsvoll, weil die Lehrkraft zweifellos einen entscheidenden Umweltfaktor für die Mehrheit der Schüler darstellt. Aufgrund der „Person-Umwelt-Transaktion“ (LEWIN) muß deshalb angenommen werden, daß sich diese Spannung, in der sich die Lehrkraft befindet, auch auf die Mehrheit der Klasse auswirken kann. Wahrnehmungsunterschiede seitens der Schüler können somit nicht nur für die Lehrkraft, sondern wiederum auch für die Mehrheit in der Klasse verhaltensrelevant werden. Durch das Ausklammern von Unterschieden in der Klimawahrnehmung bliebe somit ein potentiell wichtiger Einflußfaktor auf diesen Wechselwirkungsprozeß im dunkeln.

Während der Einfluß von solchen Wahrnehmungsunterschieden auf den Lehrer unmittelbar einleuchtet und über die Person-Umwelt-Transaktion indirekt wieder auf die meisten in der Klasse zurückgekoppelt wird, erscheint auch ein *direkter* Einfluß auf die Mehrheit der *Schüler* plausibel. In einem Unterrichtsklima, das bezüglich der subjektiven Klimawahrnehmungen durch bewußten oder unbewußten Dissens gekennzeichnet ist, kann nämlich angenommen werden, daß das Verhalten der meisten Schüler insofern beeinflusst wird, als nahezu jeder in seinem Verhalten oder Handeln (sofern es z.B. die Beziehung zur Lehrkraft betrifft) sowohl mit Zustimmung als auch mit Ablehnung seitens der Mitschüler rechnen muß. Diese Annahme gilt um so mehr, je bewußter diese Unterschiede für den Handelnden sind und je sympathischer bzw. unsympathischer die Mitschüler empfunden werden, die die Beziehung zur Lehrkraft anders einschätzen. So ist es durchaus vorstellbar, daß z.B. diejenigen Schüler, die dem Lehrer eher kritisch begegnen, jene (als unsympathisch empfundene) Mitschüler für 'Schleimer' halten, die ihn aus vermeintlich opportunistischen Gründen wohlwollend behandeln. PEKRUN weist in diesem Zusammenhang ebenfalls auf die Probleme hin, die die Einschränkung des Klimabegriffs auf die kollektive Klimawahrnehmung nach sich ziehen kann: „Eine solche Begriffseinschränkung ist u.a. deshalb problematisch, weil es die individuellen Umweltwahrnehmungen des einzelnen sind, die für ihn psychische Realität besitzen und damit direkt auf sein Verhalten einwirken ...“ (1985, S. 524).

Zu Beginn der Untersuchungen wurde angenommen, daß es sich bei den dargestellten - zum Teil sehr individuellen - Klimaausprägungen um typische Phänomene aus dem Bereich der Berufsschule handelt. Denn Berufsschulklassen sind zum einen in der Regel durch ein hohes Maß an Heterogenität in der Zusammensetzung gekennzeichnet (z.B. hinsichtlich schulischer Vorbildung, Alter, Leistungsfähigkeit, betrieblicher Herkunft, aber auch bedingt durch Gruppenbildungen aus schulorganisatorischen Gründen). Deshalb wurde vermutet, daß hier das Unterrichtsklima *individueller* erlebt wird als in meist homogener strukturierten Klassen des allgemeinbildenden Schulwesens, für das die 'Lasso' validiert wurden. Zum anderen spielt innerhalb des dualen Ausbildungssystems die Berufsschulklasse im Vergleich zum Betrieb, in dem der größte Teil der Ausbildungszeit verbracht wird, als Sozialisationsfaktor eher eine untergeordnete Rolle. Der gruppenspezifische Prozeß, der nach DREESMANN zu gemeinsam geteilten Wahrnehmungen führt, ist in Berufsschulklassen deshalb in der Regel eher schwach ausgeprägt, weshalb ebenfalls mit heterogenen Klimawahrnehmungen zu rechnen ist. Gerade in Berufsschulklassen könnten deshalb aufgrund einer zu engen Definition wichtige Informationen über die Deutung von inhaltlich bedeutsamen Klimaeinzelmerkmalen aus Sicht der Schüler verloren

gehen, die insbesondere bei bimodalen Verteilungen möglicherweise das Entscheidende an der Unterrichtsatmosphäre unberücksichtigt lassen.

DREESMANN würde bimodale Verteilungen vermutlich als Zwischenstadium beim Zusammenwachsen der Schüler zu einer Gruppe ansehen, da er von einer Homogenisierung der Schülerurteile im Laufe eines Schuljahres ausging (1982, S. 52). Was aber, wenn selbst nach zwei gemeinsamen Berufsschuljahren keine solche Homogenisierung eintritt oder sich ein anfänglich vorhandener Konsens wieder auflöst? Nachdem v. SALDERN in seinen Untersuchungen an Hauptschulen (7. Klasse) zur Frage „Homogenisierung oder Heterogenisierung der Schülerurteile?“ feststellte, daß es innerhalb eines Schuljahres nicht zu einer Homogenisierung der Urteile gekommen war (1987, S. 311), wurde für die hier vorzustellende Untersuchung entschieden, den DREESMANNschen Klimabegriff nicht zu übernehmen. Dieser erschien insbesondere für die Klimadiagnose von Berufsschulklassen, aber möglicherweise auch generell als *zu eng*. Der Klimabegriff wurde deshalb für die theoretische Verankerung der durchgeführten Untersuchungen so definiert:

Die Gesamtheit aller Gemeinsamkeiten und individuellen Unterschiede in den subjektiven Wahrnehmungen zur typischen Unterrichtsatmosphäre wird als 'Unterrichtsklima' bezeichnet.

Gemeinsame Wahrnehmungen, die bei mindestens zwei Drittel der Schüler einer Klasse übereinstimmen, werden 'Mehrheitswahrnehmung' genannt.

#### Unterrichtsklima als Gesamtheit gemeinsamer und individueller Wahrnehmungen

Der besondere Vorzug dieser Klimadefinition, die gegenüber der von DREESMANN um die *individuellen Wahrnehmungen* erweitert wurde, wird darin gesehen, daß das Fehlen von gemeinsam geteilten Klimawahrnehmungen in einer Klasse nicht mehr als Indiz für ein „nicht-existentes Klima“ (wie bei DREESMANN), sondern lediglich für eine nicht vorhandene Mehrheitswahrnehmung gilt. Der potentielle Einfluß individueller Unterschiede auf das Verhalten und Handeln bleibt damit bei einer Diagnose nicht mehr länger ausgeklammert. Insbesondere im Fall von Häufungspunkten auf unterschiedlichen Polen der Bewertungsskala stehen mit diesem erweiterten Klimabegriff möglicherweise die entscheidenden Informationen für das Verständnis des zu diagnostizierenden Unterrichtsklimas und vor allem auch für die Suche nach Verbesserungsmaßnahmen zur Verfügung.

Um die Bedeutung gemeinsam geteilter Klimawahrnehmung im Sinne DREESMANNs nicht aus den Augen zu verlieren, wurde in der Klimadefinition mit dem Begriff „*Mehrheitswahrnehmung*“ der kollektive Wahrnehmungsanteil im Sinne eines mehrheitlich geteilten Wahrnehmungskonsenses mitberücksichtigt. Denn dieser Konsens stellt (zusätzlich zu den Wahrnehmungsunterschieden) ohne Zweifel ein wichtiges Beschreibungsmerkmal für das Unterrichtsklima dar, wie es z.B. schon die Alltagsbegriffe 'Team- bzw. Gruppendeist', 'Lagerfeueratmosphäre' oder 'Volksseele' zum Ausdruck bringen. Um einen exakten Gebrauch dieses Begriffs zu ermöglichen, wurde die Zweidrittelmehrheit als Untergrenze festgelegt. Entscheidend daran ist, daß jetzt die „*Mehrheitswahrnehmung*“ von Mehrheitsverhältnissen abhängig ist und nicht mehr der Klimabegriff an sich.

Im Unterschied zu v. SALDERN wurde für die genauere Bestimmung des Klimabegriffs statt des Bestimmungswortes 'Sozial' der Begriff '*Unterricht*' gewählt, weil damit die lehrende und lernende Auseinandersetzung mit dem 'Unterrichtsgegenstand' - in Anlehnung an das didaktische Dreieck (Lehrer, Schüler, Unterrichtsgegenstand; vgl. z.B. GLÖCKEL 1990, S. 55) - stärker ins Blickfeld rückt. Dies wurde deshalb für wichtig gehalten, weil insbesondere an der Berufsschule der Unterrichtsgegenstand naturgemäß stärker im Mittelpunkt steht, als im allgemeinbildenden Schulwesen. Daß es bei Klimauntersuchungen vor allem um den emotionalen Faktor der mit der Unterrichtssituation verbundenen sozialen Prozesse geht, kommt im Grundwort 'Klima' bereits zum Ausdruck, so daß darauf m.E. nicht extra hingewiesen werden muß (wenn nicht die Gefahr einer Verwechslung mit dem meteorologischen Begriff besteht). Außerdem wird mit der Komposition 'Unterrichtsklima' das System genauer bestimmt, in dem z.B. die Analyse oder Modifikation vorgenommen wird. Vor allem weil es sich hier um ein offenes System<sup>2</sup> handelt, das von außen (Lehrerkollegium, Direktorat, Eltern, Betriebe ...) beeinflussbar ist, bleibt damit zumindest die Systemebene klar, auf die sich die Untersuchungen beziehen.

Welche weiteren Vorteile sich aus dieser Neudefinition für die Entwicklung eines Erhebungsinstrumentes sowie für die veranschaulichende Darstellung der Meßergebnisse im Fall eines Datenfeedbacks an die Betroffenen selbst ergeben, zeigt das nächste Kapitel.

---

<sup>2</sup> Vgl. SPANHEL (1993).

### NEUKONSTRUKTION DER 'LASSO' ZUR ERFASSUNG VON KLIMA- WAHRNEHMUNGEN MIT SUBJEKTIV-INDIVIDUELLEM WAHRNEHMUNGSBEZUG

Für die inhaltliche Beschreibung des Unterrichtsklimas stellten die 17 Lasso-Dimensionen (vgl. Tabelle 2.1 bis 2.3) zu Beginn der Untersuchungen den neuesten Forschungsstand in der Sozialklimaforschung dar. Auch wegen ihrer triadischen Faktorenstruktur ('Schüler-Lehrer-Beziehungen', 'Merkmale des Unterrichts', 'Schüler-Schüler-Beziehungen'), die an das bereits erwähnte didaktische Dreieck erinnert und vermutlich gerade deshalb bei Lehrkräften wie auch bei Schülern im Rahmen der Voruntersuchungen Anklang fand, wurde ihnen vor anderen Fragebögen, die ebenfalls auf Eignung für die Berufsschule hin überprüft wurden<sup>3</sup>, der Vorzug für die geplanten Klimauntersuchungen gegeben. Außerdem wurde angestrebt, kein gänzlich neues Instrument zu kreieren. Denn dies hätte sowohl die Überschaubarkeit als auch die Vergleichbarkeit der ohnehin schon großen Zahl an Klima-Erfassungsinstrumenten einschließlich der zugehörigen Forschungsergebnisse nur unnötig erschwert.

Nachdem in den 'Lasso' aber Items verwendet wurden, die bereits den kollektiven Wahrnehmungsanteil erfassen sollten, konnten sie auf Itemebene aus den genannten Gründen nicht verwendet werden. Alle 17 Skalen wurden deshalb auf der Basis theoretischer Überlegungen zur Formulierung von Klimaitems völlig neu operationalisiert. Die wichtigsten Veränderungen, die zu entscheidenden Unterschieden im Vergleich zu den 'Lasso' führten, werden in den folgenden Unterkapiteln vorgestellt.

#### Theoretische Überlegungen zur Itemformulierung

Bei der Neukonstruktion der 'Lasso' handelt es sich insofern um eine Weiterentwicklung, als versucht wurde, die inhaltliche Bedeutung der Skalen beizubehalten.

Für die anstehende Operationalisierung der 17 Skalen hatte es sich als besonders fruchtbar erwiesen, die Berufsschüler selbst bei der Item-Gewinnung zu beteiligen. Dies wurde durch entsprechende Arbeitsaufträge realisiert, bei denen die Schüler in Gruppenarbeit Formulierungen für die Beschreibung der einzelnen Klima-Subdimensionen finden sollten. Um die inhaltliche Bedeutung der Lasso-Skalen zu wahren, wurden den Schülern bei den Arbeitsaufträgen pro

---

<sup>3</sup> Z.B. 'Lei' von ANDERSON & WALBERG 1974, 'Lst' von KAHL 1977, 'Fuk' von DRÉESMANN 1982, 'Uks 84' von STUCKLE 1982.

Skala jeweils die trennschärfsten Lasso-Items als Ausgangsbeispiel vorgegeben. Auch wenn auf diese Weise natürlich noch nicht die endgültigen Itemformulierungen zur Verfügung standen, konnten dennoch wertvolle Anregungen in einer schülergerechten und dem Erleben der Schüler entsprechenden Sprache gewonnen werden, was sich positiv auf die Inhaltsvalidität auswirken sollte.

Die inhaltliche Äquivalenz zu den 'Lasso' konnte auch dadurch erzielt werden, daß mehrere Lasso-Items weiter verwendet wurden, sofern sie zu dem hier gewählten Ansatz kompatibel waren. Dies war dann möglich, wenn die Hauptvoraussetzung für die Übernahme von Items in das neu konstruierte Instrument erfüllt werden konnte, nämlich die Erfassung von *subjektiv-individuellen* Klimawahrnehmungen. Denn damit sollte die Beantwortbarkeit der Items sowie der Erfolg der Erhebung insgesamt sichergestellt werden. Einige (in inhaltlicher Hinsicht wertvolle) Lasso-Items mußten deshalb so umformuliert werden, daß sie dieser Hauptforderung genügen konnten.<sup>4</sup>

Erfahrungsgemäß wird der Begriff subjektiv-individuelle Klimawahrnehmung in der einschlägigen Literatur nicht immer mit gleichem Bedeutungsinhalt verwendet. Aus diesem Grund soll er mit Hilfe von Abbildung 1, die eine Modifikation des Klimamodells von PEKRUN (1985, S. 529) hinsichtlich des Aspekts der „Umweltwahrnehmung“ darstellt, kurz präzisiert werden.

Die Abbildung 1 soll hier vor allem zeigen, daß

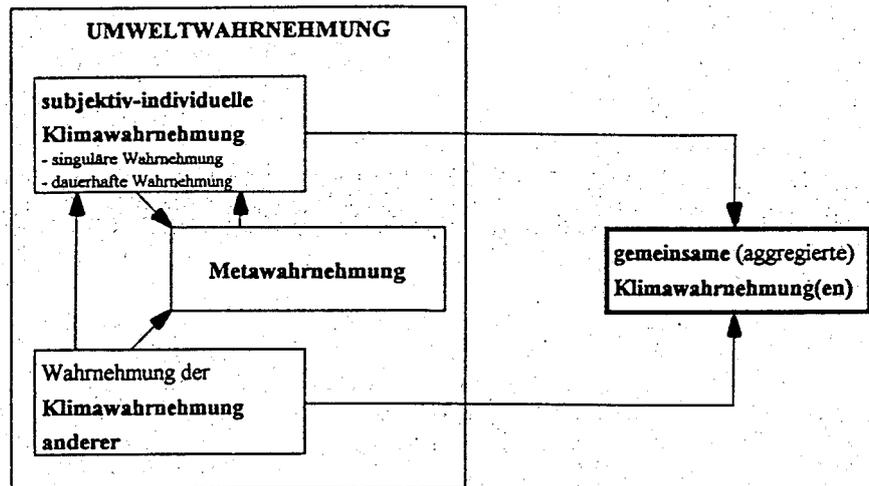
- a) die Unterrichtsklimawahrnehmung aus Sicht der Schüler bei einem subjektivistischen Zugang theoretisch in Form der „subjektiv-individuellen Klimawahrnehmung“ aller Einzelschüler oder als „Wahrnehmung der Klimawahrnehmung anderer“ (oder in Form von beiden) erfassbar ist.
- b) man unter „gemeinsamen Klimawahrnehmungen“ sowohl aggregierte subjektiv-individuelle Klimawahrnehmungen als auch wahrgenommene Klimawahrnehmungen anderer verstehen kann, die ebenfalls (z.B. auf Klassenebene) aggregierbar sind.

---

<sup>4</sup> Insgesamt sind im neu konstruierten Instrument, das 126 Items umfaßt, noch 37 Lasso-Items enthalten, die teils im Original, teils mit geringen sprachlichen Veränderungen eingegangen sind (vgl. MÜLLER 1996, Anhang 6f).

Abbildung 1

Modell zur Umweltwahrnehmung einer Person nach PEKRUN (modifiziert)



Zu a)

Während mit den von Berufsschülern vorgeschlagenen Items die subjektiv-individuelle Klimawahrnehmung gemessen würde, erfaßt z.B. das Lasso-Item 87 die subjektive „Wahrnehmung der Klimawahrnehmung anderer“. Damit kann bei einer subjektiven Klimabeurteilung im Sinne eines sozialperzeptiven Ansatzes sowohl ein individueller als auch ein kollektiver Bezug vorliegen, der im folgenden als 'Wahrnehmungsbezug' bezeichnet wird.

Bezugnehmend auf SIROTNIK (1980), wurden bei den 'Lasso' drei Typen von Items verwendet, die auf den ersten Blick genau diese Unterscheidung von Wahrnehmungen widerspiegeln: „Einmal die personenzentrierten sog. 'Ich-Items' (= subjektiv individueller Bezug), dann die gruppencentrierten 'Sie-Items' (= subjektiv kollektivistischer Bezug) und schließlich die Items, die qualitativ dazwischen stehen, die 'Wir-Items', welche den Probanden in die Gruppe einbeziehen“ (v. SALDERN 1987, S. 222). Zudem wurden noch Items unterschieden, die sich „... allgemein auf die Umwelt beziehen, ohne daß der Antwortende als Bezugspunkt genannt wird, z.B. 'Unser Klassenzimmer ist geräumig'.“ (v. SALDERN, a.a.O.).

Im letzten Beispiel wird der Wahrnehmungsbezug zwar nicht explizit genannt, dennoch wird der Antwortende diese Aussage natürlich auf sich selbst, also auf seine subjektiv-individuelle Wahrnehmung beziehen. Inhaltlich ergäbe sich ja auch kein Unterschied, wenn das Item so formuliert wäre: Ich finde unser Klassenzimmer geräumig. Warum sollte auch jemand auf die Idee kommen, das Item so zu verstehen: Die Mehrheit in der Klasse findet unser Klassenzimmer geräumig? Obwohl dieses Item in der Lasso-Konzeption nicht als „Ich-Item“ bezeichnet wird (in diesem Fall hätte es ja auch ausgeschlossen werden müssen!), hat es insofern die Funktion eines solchen, als ein subjektiv-individueller Wahrnehmungsbezug zumindest implizit vorliegt, bzw. vom Probanden mit großer Wahrscheinlichkeit so aufgefaßt würde. Analog gilt dies z.B. auch für das Item 6.9 („In diesem Fach wird viel verlangt“) aus dem neu konstruierten Instrument, mit dem u.a. der „Leistungsdruck“ als Merkmal des Unterrichts erfaßt wird.

Weder das Fehlen des Personalpronomens 'ich', noch das Vorhandensein des Pronomes 'unser' oder die Verwendung von adverbialen Bestimmungen (z.B. 'viele', 'die meisten' o.ä.) sagt also per se etwas darüber aus, ob sich der Beantworter eines Fragebogens auf seine eigenen Klimawahrnehmungen oder auf die wahrgenommene Klimawahrnehmung anderer bezieht. Bei der Itemformulierung (bzw. -selektion) muß dieser Wahrnehmungsbezug also geklärt werden, und zwar unabhängig von der Verwendung bestimmter Personalpronomina oder adverbialer Bestimmungen, da sie nicht unbedingt etwas über den Wahrnehmungsbezug aussagen. Aus diesen Gründen wurden auch in den neu konstruierten Skalen nicht ausschließlich 'Ich-Items' verwendet, wie vermutet werden könnte. Ein bewußter Verzicht auf solche Items war sogar bei manchen Skalen (z.B. „Aggressionen gegen die Lehrkraft“) nötig, um der bekannten Neigung von Probanden entgegenzuwirken, im Sinne *sozialer Erwünschtheit* zu antworten.

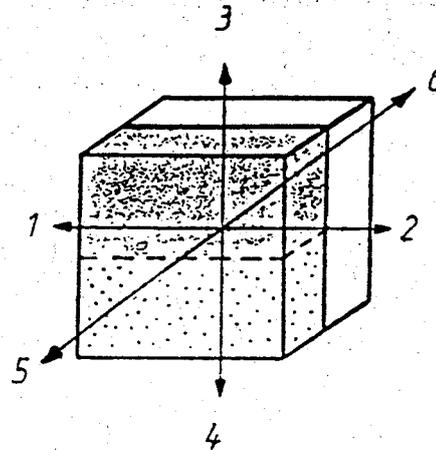
Aufgrund der vielen Fragen bezüglich der richtigen Itemformulierung und Eignung von Lasso-Items für das neu konstruierte Instrument, konnte der dimensionsanalytische Ansatz zur Kategorisierung von Itemtypen von DANN, HELMREICH, MÜLLER-FOHRBRODT, PFEIFFER und CLOETTA (1975) eine wertvolle Hilfe sein. Dort werden die Itemtypen nach zwei Kategorien eingeteilt: nach 'Objektbezug' (eigenes versus fremdes Verhalten) und nach 'Interpretationsspielraum' (Beschreibung versus Interpretation des Verhaltens). Nachdem sich letztlich herausstellte, daß das Hauptproblem der mißglückten Erhebung mit dem 'Lasso' vor allem im 'Wahrnehmungsbezug' gelegen hatte, wird eine Erweiterung dieses zweidimensionalen Ansatzes um eine dritte Dimension vorgeschla-

gen, um diesen Wahrnehmungsbezug bei zukünftigen Konstruktionen von Klimaskalen nicht mehr aus den Augen zu verlieren.

Wie in Abbildung 2 gezeigt, wurde der 'Wahrnehmungsbezug' durch die beiden Pole „subjektiv-individuelle Klimawahrnehmung“ auf der einen und subjektiv-kollektivistische Wahrnehmung („Wahrnehmung der Klimawahrnehmung anderer“) auf der anderen Seite definiert.<sup>5</sup>

Abbildung 2

*Kategoriensystem für Klimaitems nach DANN et al. (modifiziert)*



<sup>5</sup> Da aufgrund des (weit verbreiteten) Klimabegriffs von DREESMANN die z.B. von PEKRUN (1985, S. 530) gewählte Bezeichnung „individuelles Klima“ als Widerspruch und „kollektives Klima“ als Tautologie aufgefaßt werden könnte, wurden diese Bezeichnungen hier vermieden. Statt dessen wurde stets von Klimawahrnehmung gesprochen. Nachdem sich die Begriffe 'individuelles' und 'kollektives Klima' auf die Unterscheidung des „beta press“ in „individual und consensual beta press“ (STERN 1970) zurückführen lassen, ist in beiden Fällen die subjektive Wahrnehmung des Klimas ('beta press') gemeint. In einem subjektivistischen Ansatz ist Klima demnach gleichbedeutend mit Klimawahrnehmung. So gesehen handelt es sich also bei den Bezeichnungen 'individuelles und kollektives Klima' gleichzeitig um die individuelle und kollektive Wahrnehmung der Unterrichts-atmosphäre aus der subjektiven Sicht der Befragten. Deshalb besteht zwischen den hier gewählten Begriffen und denen von PEKRUN kein wirklicher Unterschied.

Objektbezug	Interpretationsspielraum	Wahrnehmungsbezug
1 = fremdes Verhalten	3 = Beschreibung	5 = subjektiv-individuell
2 = eigenes Verhalten	4 = Interpretation	6 = subjektiv-kollektivistisch

*Erläuterungen:* Der gepunktet dargestellte Bereich innerhalb des Würfels kennzeichnet die Items, die bei der Konstruktion des vorgelegten Meßinstruments verwendet wurden. Dieser Bereich soll verdeutlichen, daß konsequent darauf geachtet wurde, nur Items mit subjektiv-individuellem Wahrnehmungsbezug zu verwenden. Beispielitem 2.8: „Der Lehrer würde mir bestimmt helfen, wenn ich mit einem Problem zu ihm käme“.

Der stärker gepunktete obere Bereich soll zum Ausdruck bringen, daß versucht wurde, den Schwerpunkt auf Items mit geringem Interpretationsspielraum zu legen, sofern dies die jeweilige Klimadimension zuließ.

Beispielitem 12.7: „Die Klasse zerfällt in verschiedene Cliquen, die kaum zusammenarbeiten.“

#### Zu b)

Der subjektiv-individuelle Wahrnehmungsbezug bei der Klimaerfassung bietet die Möglichkeit, im Sinne einer Mehrebenenanalyse auf unterschiedlichen Aggregatenebenen Gemeinsamkeiten und Unterschiede in den einzelnen Wahrnehmungen rechnerisch zu ermitteln. Damit sind zuverlässige Klimadiagnosen möglich, weil sie nicht auf Mutmaßungen darüber beruhen, wie z.B. die Mehrheit der Klasse die einzelnen Klimabereiche *wahrscheinlich* bewerten würde (sofern die Schüler überhaupt zu solchen Angaben in der Lage sind).

Freilich muß auch bei diesem Klimaansatz offen bleiben, welche Beziehung zwischen dem vorgestellten erweiterten Klimabegriff und überindividuellen Wissenssystemen besteht, die sich als ganzheitliche Qualitäten nicht ohne weiteres aus Einzelkognitionen errechnen lassen.<sup>6</sup> Wie noch zu zeigen sein wird, wurde hier jedoch angenommen, daß durch Berücksichtigung von 'Klassenmittelwert' und 'Häufigkeitsverteilung' subjektiv-individueller Klimawahrnehmungen das *überindividuelle* Phänomen adäquater beschrieben werden konnte als in bisherigen Konzeptionen mit der unscharfen Festlegung „die meisten in einer Klasse“. Vor allem wurde aber zum hier gewählten Ansatz aufgrund der dargestellten Beantwortungsproblematik von Fragebogenitems mit subjektiv-kollektivistischem Wahrnehmungsbezug keine überzeugende Alternative gesehen.

<sup>6</sup> Im wesentlichen wird damit das sog. 'Emergenzproblem' der Philosophie tangiert, wie es z.B. in folgender Aussage aus der Gestaltpsychologie zum Ausdruck kommt: „Das Ganze ist mehr als die Summe seiner Teile“ (z.B. NAGEL, 1967).

*Insgesamt* besteht (neben einigen sprachlichen Veränderungen bei den Skalenbezeichnungen) das Neue an dem hier entwickelten Instrument also darin, daß sich die Schüler bei der Beantwortung der Items auf ihre subjektiv-individuelle Wahrnehmung beziehen können. Auf der Erfassungsebene werden somit subjektiv-individuelle Klimawahrnehmungen erhoben, um daraus auf der Auswertungsebene die Gemeinsamkeiten und Unterschiede (z.B. auf Klassenebene) rechnerisch zu ermitteln.

#### Kurzdarstellung der neu konstruierten Unterrichtsklima-Skalen

In diesem Kapitel soll nun das neu konstruierte Instrument näher vorgestellt werden, wie es in der Hauptuntersuchung verwendet wurde.

#### Grobübersicht über das Skalenkonzept

Die Tabellen 2.1 bis 2.3 zeigen die 17 Einzeldimensionen mit dem jeweils trennschärfsten Item. Außerdem wurde die interne Konsistenz der Skalen (Cronbachs Alpha) sowie die mittlere Trennschärfe des Beispielitems angegeben, wie sie für den 1. Meßzeitpunkt ermittelt wurden. An der 1. Messung nahmen 41 Klassen mit insgesamt 852 Schülern unterschiedlicher Berufsfelder teil. Das gesamte Instrument umfaßte insgesamt 126 Items.

Tabelle 2.1

#### Schüler-Lehrer-Beziehungen

1.	Schüler-Lehrer-Beziehungen	Itemzahl	Cronbachs $\alpha$	Trennschärfe
1.	Fürsorglichkeit der Lehrkraft <i>1.6: Der Lehrer würde mir bestimmt helfen, wenn ich mit einem Problem zu ihm käme.</i>	7	.81	.61
2.	Aggressionen gegen die Lehrkraft <i>2.8: Der Lehrer wird oft absichtlich geärgert.</i>	9	.89	.75
3.	Bevorzugung/Benachteiligung <i>3.11: Ich habe das Gefühl, der Lehrer mag mich nicht.</i>	8	.87	.68
4.	Autoritäre Merkmale im Unterrichtsstil <i>4.9: Die Lehrkraft läßt uns oft spüren, daß sie (und nicht die Mehrheit der Klasse) das Sagen hat.</i>	8	.81	.60
5.	Zufriedenheit mit der Lehrkraft <i>5.7: Der Lehrer ist mir sympathisch.</i>	7	.86	.72

Tabelle 2.2

*Merkmale des Unterrichts*

II.	Merkmale des Unterrichts	Itemzahl	Cronbachs $\alpha$	Trennschärfe
6.	Leistungsdruck 6.8: <i>In diesem Fach fühle ich mich oft überfordert.</i>	8	.88	.72
7.	Disziplin/Ordnung in der Klasse 7.4: <i>Unser Lehrer müsste mehr für Ruhe und Ordnung während des Unterrichts sorgen.</i>	10	.89	.79
8.	Fähigkeiten zur Stoffvermittlung 8.12: <i>Der Unterrichtsaufbau des Lehrers ist meist klar und logisch.</i>	9	.90	.78
9.	Mitarbeit der Schüler 9.10: <i>In diesem Fach bin ich häufig nicht richtig bei der Sache.</i>	9	.84	.65
10.	Resignation bei den Schülern 10.4: <i>In diesem Unterrichtsfach habe ich kaum Aussicht auf Erfolg, egal wie sehr ich mich anstreng.</i>	5	.90	.82
11.	Zufriedenheit mit dem Unterricht 11.1: <i>Mir macht dieses Unterrichtsfach keinen Spaß.</i>	6	.83	.70

Tabelle 2.3

*Schüler-Schüler-Beziehungen*

III.	Schüler-Schüler-Beziehungen	Itemzahl	Cronbachs $\alpha$	Trennschärfe
12.	Cliquenbildung unter den Schülern 12.7: <i>Die Klasse zerfällt in verschiedene Cliquen, die kaum zusammenarbeiten.</i>	5	.87	.76
13.	Hilfsbereitschaft unter den Schülern 13.7: <i>Wenn ich bei der Vorbereitung z.B. auf Schulaufgaben Probleme hätte, würden mir (vermutlich) die meisten in der Klasse helfen.</i>	7	.87	.68
14.	Aggressionen zwischen den Schülern 14.5: <i>Unter den Mitschülern herrscht oft eine aggressive Stimmung.</i>	7	.91	.77
15.	Diskriminierung von Schülern 15.5: <i>Bestimmte Mitschüler erfahren wenig Anerkennung von der Klasse.</i>	9	.84	.60
16.	Konkurrenzdenken 16.7: <i>In unserer Klasse herrscht ein ständiger Notenkampf.</i>	6	.85	.69
17.	Zufriedenheit mit den Mitschülern 17.8: <i>Meine Mitschüler sind fast alle nett.</i>	6	.84	.68

*Gütekriterien*

Nach der Operationalisierung und Präzisierung der Skalenkonzepte wurden die neu konstruierten Skalen zunächst im Rahmen von Voruntersuchungen bei zwei Analytestichproben eingesetzt. Die nach den Itemselektionen durchgeführten Reanalysen erbrachten bereits akzeptable Reliabilitätswerte. In der Kreuzanalyse mit den Daten der ersten Messung der Hauptuntersuchung ergaben sich für die interne Konsistenz der Skalen Werte für Cronbachs Alpha zwischen .81 und .91 (MÜLLER 1996, S. 164).

Sowohl in den Vortests als auch in der Hauptuntersuchung zeigte sich am Echo der an die Klassen zurückgemeldeten Meßergebnisse, daß mit den neu operationalisierten Skalen ein Meßinstrument geschaffen werden konnte, dessen Meßdaten von den Betroffenen selbst, d.h. jeweils von Lehrkraft und Klasse, durchweg akzeptiert wurden. Angesichts der Beteiligung von Berufsschülern bei der Suche nach geeigneten Itemformulierungen ist dies auch nicht verwunderlich, denn dadurch dürfte ein entscheidender Beitrag für die inhaltliche Validität geleistet worden sein. Dies gilt vor allem auch in Verbindung mit den vorgegebenen Lasso-Skalenbezeichnungen, da diese ja selbst aus einer umfangreichen Forschungsarbeit als valide Merkmale zur Beschreibung des Klima-Konstrukts hervorgegangen waren.

Diese Sichtweise zur Validität konnte auch durch Faktorenanalysen bestätigt werden, bei denen sich z.B. innerhalb einer Grobdimension (z.B. Schüler-Lehrer-Beziehungen) nahezu alle Einzelitems den konstruierten Subdimensionen zuordnen ließen. Auch eine über alle Subdimensionen hinweg durchgeführte Faktorenanalyse bildete die drei Lasso-Grobdimensionen (I, II, III) gut ab, wobei hier ein *zusätzlicher* Faktor mit einer aufklärenden Varianz von 6.5% gefunden wurde, der die Skalen 2, 7, 9 und 14 (vgl. Tabelle 2.1 bis 2.3) vereinigt und als 'systemimmanenter Störfaktor' bezeichnet werden kann.

Die Konstruktvalidität konnte schließlich auch durch Skaleninterkorrelationen zwischen inhaltlich ähnlichen und verschiedenen Skalen bestätigt werden (vgl. MÜLLER, 1996, S. 148 ff)

Bevor mit den Hauptbefunden der Längsschnittuntersuchung noch weitere Hinweise auf die Validität des Instruments vorgestellt werden, soll noch auf einige meßtechnische Besonderheiten hingewiesen werden, die sich aus dem neudefinierten Klimabegriff ergeben.

Meßtechnische Besonderheiten

Für die quantitative Beschreibung von gemeinsamen und individuellen Klimawahrnehmungen im Sinne des neudefinierten Klimabegriffs wurden - wenn es um eine Rückmeldung der Meßergebnisse an Lehrkraft und Klasse ging - der Klassenmittelwert und die Häufigkeitsverteilung verwendet.<sup>7</sup>

Während der arithmetische *Klassenmittelwert* das durchschnittliche Wahrnehmungsurteil der Schüler auf Klassenebene darstellt, wurde für die Beschreibung der individuellen Unterschiede in den Klimawahrnehmungen die *Häufigkeitsverteilung* aller individuellen Klimaurteile berechnet. Beides wurde pro Subdimension auf zwei Farbfolien veranschaulicht. Einen Ausschnitt aus dieser Folie für eine Gesamtauswertung einer Klasse, die an der Hauptuntersuchung mit drei Messungen innerhalb eines Schuljahres teilnahm, zeigt Tabelle 3.

Tabelle 3

Ausschnitt aus einer Meßwertanzeige für die Gesamtauswertung von Klimaerhebungen in einer Klasse über drei Meßzeitpunkte

	negativ					positiv				
	-4	-3	-2	-1	0	+1	+2	+3	+4	
	wenig ...					...viel				
Fürsorglichkeit der Lehrkraft:	1				2	14	4	1		
	2				1	5	13	1		
	3			2	14	3				
	viel ...					...wenig				
Aggressionen gegen die Lehrkraft:	1	1	2	1	5	2		7	3	
	2		2	1	5	2	4	6		
	3		2	2	5	2	4	4		

Anmerkungen: Die Abbildung zeigt die Häufigkeiten der Antwortalternativen in einer Klasse. Die grauunterlegten Kästchen zeigen die Höhe des Klassenmittelwertes bei der jeweiligen Messung. Die einzelnen Zahlen in den Kästchen verdeutlichen die Häufigkeitsverteilung, d.h. die Anzahl der Schüler, deren Skalenmittelwerte in den zugehörigen Bewertungsintervallen lagen. Linke Spalte: 1 = erste Messung, 2 = zweite Messung, 3 = dritte Messung.

Auch hier wird bei Merkmal 2 wieder zu fast allen Meßzeitpunkten eine bimodale Verteilung deutlich (wenn man alle positiven bzw. negativen Urteile

<sup>7</sup> Bei der theoretischen Analyse der Unterschiede in den Wahrnehmungen wurde statt der Häufigkeitsverteilung die *Standardabweichung* als Streuungsmaß verwendet.

zusammenfaßt)<sup>8</sup>. Dies war auch bei anderen Klimamerkmale in weiteren Klassen der Hauptuntersuchung festzustellen.

Im letzten Kapitel soll nun noch gezeigt werden, wie mit Hilfe des entwickelten Instruments im Zuge der Längsschnittstudie zur Wirkungsweise einer Survey-Feedback-Methode durchaus plausible Forschungsergebnisse erzielt werden konnten.

#### LÄNGSSCHNITTSTUDIE ZUR KLIMAVERBESSERUNG UNTER EINSATZ DER NEUEN UNTERRICHTSKLIMA-SKALEN ALS SURVEY-FEEDBACK-INSTRUMENT

Bei der Längsschnitt-Untersuchung ging es um den Versuch, das Unterrichtsklima durch Erfassung der Schülersicht ('survey') und anschließende Rückmeldung der Meßergebnisse an Lehrkraft und Klasse ('feedback') zu verbessern. Nachdem mit 'survey-feedback'-Methoden sowohl bei der Organisationsentwicklung von Betrieben (vgl. z.B. TÖFFER, ZANDER & BECK, 1985) als auch im Bereich der Grundschule (vgl. DUTKA, 1985) bereits Klimaverbesserungen bewirkt werden konnten, sollte nun überprüft werden, inwieweit dies auch an der Berufsschule möglich ist.

#### Kurze Versuchsbeschreibung

Die Treatmentstudie wurde im Schuljahr 1992/93 an mittelfränkischen Berufsschulen durchgeführt. Von den 41 teilnehmenden Klassen konnten für die Auswertung der drei durchgeführten Messungen die Daten von 35 Klassen mit insgesamt 660 Schülern, die an allen drei Messungen teilnahmen, genutzt werden. An der Untersuchung nahmen Schüler unterschiedlicher Jahrgangsstufen, Schulungsformen und Berufsfelder zusammen mit ihrem Klassenleiter teil, so daß insgesamt von einer ausreichenden Repräsentativität ausgegangen werden kann. Die Überprüfung der Ergebnisse erfolgte mittels eines randomisierten Zwei-Gruppen-Versuchsplans mit zwei wiederholten Messungen im Abstand von jeweils drei Monaten (vgl. Tabelle 4).

---

<sup>8</sup> Es handelte sich hier um eine (Vergleichs-)Klasse des 12ten Jahrgangs, bei der sich Schüler und Lehrkraft mehr als zwei Jahre lang kannten.

Tabelle 4

*Versuchsdesign für die Längsschnittuntersuchung*

		Messung 1	Treatment-phase 1	Messung 2	Treatment-phase 2	Messung 3
R	Vergleichsgruppe	Okt./Nov. 92	—	Febr./März 93	—	Mai/Juni 93
	Testgruppe	Okt./Nov. 92	Nov. 92-Febr. 93	Febr./März 93	März 93-Mai 93	Mai/Juni 93

Während in der Testgruppe nach der ersten, zweiten und dritten Messung eine standardisierte Rückmeldung der Ergebnisse erfolgte, fand in der Vergleichsgruppe eine Rückmeldung erst nach der dritten Messung statt.

Das Treatment bestand also darin, daß die von mir errechneten Meßergebnisse an die jeweilige Lehrkraft der Testgruppe weitergereicht wurden, die diese mit ihrer Klasse besprechen sollten. Außerdem sollten sie gemeinsam mit den Schülern ggf. nach Verbesserungsmöglichkeiten suchen. Die dabei getroffenen Abmachungen wurden sowohl von der Lehrkraft als auch von den Klassensprechern protokolliert und überwacht.

Um sicherzustellen, daß erwartete Verbesserungseffekte in den Testklassen allein auf das Datenfeedback zurückzuführen sind, wurde - zusätzlich zur Randomisierung - den Teilnehmern gegenüber nicht erwähnt, daß es um die Überprüfung der Wirksamkeit einer Survey-Feedback-Methode ging. Statt dessen erhielten alle Teilnehmer in den Versuchsunterlagen die Information, es würde hauptsächlich um die Erforschung von Wahrnehmungsunterschieden in den einzelnen Klimadimensionen zwischen Klasse und Lehrkraft gehen, was in Wirklichkeit nur von nebensächlicher Bedeutung war (cover story).

Für die Auswertung des komplexen Datenmaterials wurde im Sinne eines hierarchischen Prüfplans eine Kombination aus konfirmatorischen und exploratorischen Analysen durchgeführt, um aufgrund der Vielzahl von Einzeltests Zufallssignifikanzen zu vermeiden und die Ergebnisse der Signifikanzprüfungen so gut wie möglich abzusichern. Nachdem im Sinne konfirmatorischer Signifikanzprüfung lediglich für den Klimabereich I (Schüler-Lehrer-Beziehungen) signifikante Ergebnisse auf dem 5%-Niveau erzielt werden konnten, beziehen sich die folgenden Ausführungen nur noch auf den Klimabereich I.

## Wichtige Hauptbefunde

*Veränderung des durchschnittlichen Klassenmittelwertes*

## a) Ergebnis

Als gesichertes Hauptergebnis ergab sich eine positive Gesamtveränderung der Schüler-Lehrer-Beziehungen durch das Treatment, gemessen am durchschnittlichen Klassenmittelwert ( $p < .02$ ; vgl. Tabelle 5). Die Schüler-Lehrer-Beziehungen sind damit in einer durchschnittlichen Testklasse über ein Schuljahr hinweg ( $t_3-t_1$ ) um 1,7% besser und in einer durchschnittlichen Vergleichsklasse um 4,2% schlechter geworden (vgl. Abbildung 3).

## Abbildung 3

*Veränderung des durchschnittlichen Klassenmittelwertes über drei Meßzeitpunkte bei den Schüler-Lehrer-Beziehungen*

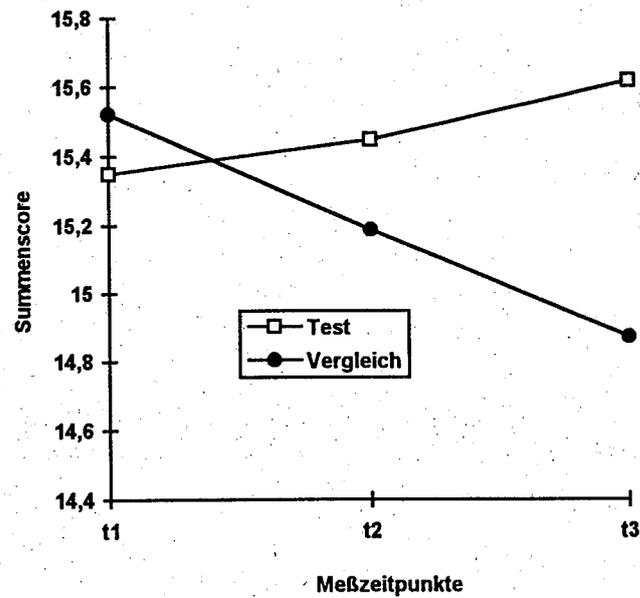


Tabelle 5  
Diagramm Daten zu Abbildung 3

	N	t2-t1	p-Wert	t3-t2	p-Wert	t3-t1	p-Wert
Vergleichsgruppe	18	-2,1 %		-2,1%		-4,2%	
			.167		.060		.013
Testgruppe	17	+0,6%		+1,1%		+1,7%	

Anmerkungen: N = Anzahl der Klassen;  $t2-t1 = (M2-M1)/M1*100$ ;  $t3-t2 = (M3-M2)/M2*100$ ;  $t3-t1 = (M3-M1)/M1*100$ , mit M = durchschnittlicher arithm. Klassenmittelwert für den Summenscore der fünf Subdimensionen der Schüler-Lehrer-Beziehungen pro Messung. Die p-Werte für die drei Veränderungswerte (t2-t1, t3-t2, t3-t1) wurden mit dem T-Test ermittelt.

#### b) Diskussion

Die positive Gesamtveränderung im Durchschnitt aller Testklassen entsprach von der Richtung her betrachtet den Erwartungen, die in den Versuchshypothesen formuliert wurden. Grundlage für die erwarteten Verbesserungen im Bereich I waren z.B. die schon erwähnten Untersuchungen von DUTKA (1985), der im Bereich der Grundschule durch Survey-Feedback-Techniken ebenfalls Klimaverbesserungen im Bereich der Schüler-Lehrer-Beziehungen herbeiführen konnte. Obwohl in bezug auf die Effektstärke - gemessen an den Ergebnissen von DUTKA - ein höheres Niveau erwartet worden war, können seine Befunde zumindest bezüglich der Veränderungsrichtung mit den neuen Skalen bestätigt werden.

Für den eher niedrig einzustufenden Verbesserungseffekt konnten auf der Grundlage der verfügbaren Daten viele Ursachen gefunden werden, wie z.B. relativ hohe Baselinewerte<sup>9</sup> oder dem Durchschnittsergebnis gegenläufige Entwicklungen in fast der Hälfte aller Einzelklassen. Vor allem darin, daß eine Klimaverbesserung zwar im Durchschnitt aller Klassen, nicht jedoch in jeder Einzelklasse möglich war, zeigen sich offenbar die Grenzen von Survey-Feedback-Methoden im Bereich der Berufsschule: Sie können einen wertvollen Impuls zur Verbesserung des Unterrichtsklimas geben, aber den Erfolg nicht in jedem Fall garantieren.

Wie die Protokollnotizen seitens der Lehrkräfte und Klassensprecher zeigen, besteht in bezug auf die Gesprächsführung beim Datenfeedback sowie bei den folgenden pädagogischen Maßnahmen zur Klimaverbesserung und bei deren

<sup>9</sup> Mehr als die Hälfte der Testklassen erzielten bei Messung 1 bereits Werte über 75% vom Skalenmaximum.

Implementierung allerdings noch Forschungsbedarf, um in Form von Modellen, Methoden und Techniken den Lehrkräften erfolgversprechende Wege zeigen zu können, auf die sie von alleine im Schulalltag nicht ohne weiteres kommen würden.

Besonders bemerkenswert ist die Entwicklung in den *Vergleichsklassen*, weil eine Verschlechterung nicht erwartet wurde. Interessant an den Verschlechterungen in den Vergleichsklassen war die Frage, ob diese der *typischen* Entwicklung im Schulalltag entspricht, oder ob es sich um die Auswirkungen des Systemeingriffs handelt, der durch die Messung in unvermeidbarer Weise entstanden ist. Auch wenn dies natürlich nicht mit letzter Sicherheit geklärt werden kann, wurde nach Abwägung aller verfügbaren Daten davon ausgegangen, daß es sich bei der negativen Entwicklung in den Vergleichsklassen mit großer Wahrscheinlichkeit eher um eine typische Entwicklung als um die Auswirkungen des Systemeingriffs handelt (auch wenn eine Überlagerung beider Effekte ebenso möglich ist). Dafür sprechen z.B. die getroffenen methodischen Vorkehrungen sowie die theoretische Annahme, daß sich die Schüler-Lehrer-Beziehungen deshalb verschlechtern können, weil Lehrkräfte im Laufe eines Schuljahres neben den problemhaltigen unterrichtlichen Aufgaben (wie z.B. Bewerten und Zensieren) auch die meist belastenden Erziehungsaufgaben (z.B. Disziplinieren und Moralisieren) bewältigen müssen, denen manche Lehrkräfte gerne ausweichen. Am wenigsten ist ein solches Ausweichen erfahrungsgemäß für den Klassenleiter einer Berufsschule möglich, was gerade für ihn schnell zu (u.U. dauerhaften) Störungen auf der Ebene der Beziehungen zu seinen Schülern führen kann. Da die Untersuchung lediglich mit Klassenleitern durchgeführt wurde, erscheint die Interpretation des Ergebnisses in den Vergleichsklassen als *typische* Entwicklung im Verlauf eines Schuljahres somit plausibel.

#### *Veränderung der durchschnittlichen Streuung*

##### a) Ergebnis

Gemessen an der Streuung der subjektiv-individuellen Schülerwahrnehmungen in den einzelnen Klassen der Testgruppe wurde im Gesamteffekt eine tendenzielle Homogenisierung erkennbar (vgl. Abbildung 4). Zwischen Messung 1 und 3 hatte sich die durchschnittliche Standardabweichung in den Vergleichsklassen um ca. 16% erhöht, während sie in den Testklassen um 4% gesunken war ( $p < .10$ ; vgl. Tabelle 6). Dieser Gesamteffekt ist offenbar auf einen signifikanten Effekt in der zweiten Treatmentphase zurückzuführen ( $p = .012$ ).

Abbildung 4

Veränderungen der Durchschnittsstreuung über drei Meßzeitpunkte bei der subjektiv-individuellen Bewertung der Schüler-Lehrer-Beziehungen

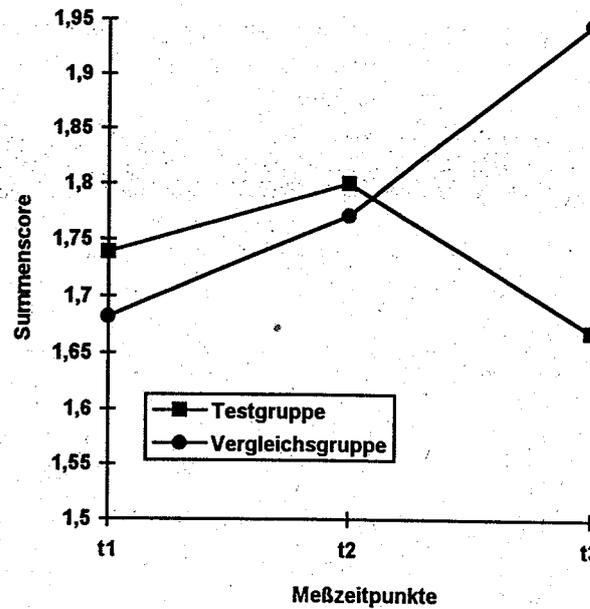


Tabelle 6

Diagrammdaten zu Abbildung 4

	N	t2-t1	p-Wert	t3-t2	p-Wert	t3-t1	p-Wert
Vergleichsgruppe	18	5,4%		9,8%		15,7%	
Testgruppe	17	3,6%	.842	-7,3%	.012	-4,0%	.092

Anmerkungen: N = Teilnehmerzahl;  $t2-t1 = (SD2-SD1)/SD1*100$ ;  $t3-t2 = (SD3-SD2)/SD2*100$ ;  $t3-t1 = (SD3-SD1)/SD1*100$ , mit SD = durchschnittliche Standardabweichung der jeweiligen Gruppe für den Summenscore der fünf Subdimensionen der Schüler-Lehrer-Beziehungen. Die p-Werte für die drei Veränderungswerte (t2-t1, t3-t2, t3-t1) wurden mit dem T-Test ermittelt.

### b) Diskussion

In den *Testklassen* konnte das erste Klassengespräch offenbar noch keine Gruppenunterschiede bezüglich einer Streuungsveränderung hervorrufen. Vermutlich wurden durch die 1. Datenrückkopplung die unterschiedlichen Urteile zur Schüler-Lehrer-Beziehung im Sinne der 'Metawahrnehmung' (vgl. Abbildung 1) zwar bewußt, konnten aber noch keinen so starken gruppendynamischen Prozeß auslösen, daß bereits bei der 2. Messung ein deutlicher Gruppenunterschied feststellbar gewesen wäre. Erst nach dem zweiten Klassengespräch kam es offensichtlich zu einer (tendenziellen) Angleichung individueller Unterschiede. Wahrscheinlich wurde durch die zweite Datenrückkopplung der Prozeß der Metawahrnehmung noch einmal angestoßen, so daß sich durch erneutes Gewährwerden von Unterschieden zwischen den eigenen subjektiv-individuellen Wahrnehmungen und den Wahrnehmungen anderer eine Homogenisierung während der zweiten Treatmentphase einstellen könnte.

Die tendenziell signifikante Zunahme überindividueller Wahrnehmung der Schüler-Lehrer-Beziehungen ist im Sinne eines größeren Konsenses in jedem Fall als positives Ergebnis zu bewerten, vor allem wenn davon ausgegangen wird, daß ein vorhandener Dissens (insbesondere bei bimodalen Verteilungen) belastende Spannungen in der Unterrichtsatmosphäre darstellen kann, je nachdem, wie weit die Häufungspunkte auseinander liegen.

In den *Vergleichsklassen* wurde (unter Berücksichtigung der bereits genannten Einschränkungen aufgrund des Systemeingriffs) die tendenziell signifikante Zunahme der durchschnittlichen Streuung im Gesamteffekt als Indiz dafür betrachtet, daß sich im Laufe eines Schuljahres eher eine zunehmende Individualisierung der Wahrnehmung der Schüler-Lehrer-Beziehungen in Berufsschulklassen vollzieht, wenn über das Unterrichtsklima im Sinne des Treatments nicht gesprochen wird. Denn dies entspricht den bereits erwähnten Ergebnissen von v. SALDERN, die zum einen die Annahme stützen können, bei den Veränderungen in den Vergleichsklassen handele es sich um eine typische Entwicklung einer durchschnittlichen Berufsschulklasse im Laufe eines Schuljahres. Zum anderen können die ähnlichen Ergebnisse wieder auf die Validität der neu konstruierten Skalen hinweisen.

Neben diesen beiden Hauptergebnissen, die auf konfirmatorischem Wege gefunden wurden und sichere Belege auf dem 5%-Signifikanzniveau darstellen, konnten bei den exploratorischen Analysen eine Reihe von weiteren interessanten Nebenergebnissen gewonnen werden. Diese erlauben einerseits einen tieferen Einblick in die Veränderungsprozesse (z.B. in anderen Klimadimensionen oder

in unterschiedliche Treatmenteffekte bei den einzelnen Jahrgangsstufen), als es im Rahmen dieses Artikels vorzustellen möglich war. Andererseits stellen sie weitere Hinweise auf die Validität des neu konstruierten Instruments und damit auf die Berechtigung des hier vorgestellten Klimakonzepts dar, bei dem - in Erweiterung zum DREESMANNschen Klimabegriff - auch Klimawahrnehmungen mit 'subjektiv-individuellem Wahrnehmungsbezug' berücksichtigt wurden.

## LITERATUR

- ANDERSON, G.J. & WALBERG, H.J. (1974). Learning environments. In H. WALBERG (Hrsg.), *Evaluating educational performance. A sourcebook of methods, instruments and examples*. Berkeley, CA: McCutchan.
- BESSOTH, R. (1989). *Verbesserung des Unterrichtsklimas*. Neuwied u. Frankfurt: Luchterhand.
- DANN, H.-D., HELMREICH, R., MÜLLER-FOHRBRODT, G., PFEIFFER, H. & CLOETTA, B. (1975). Ist Umwelt meßbar? - Ein Versuch zur Erfassung von Ausbildungsbedingungen für den innovativen Lehrer. In H. WALTER (Hrsg.), *Sozialisationsforschung, Band III: Sozialökologie - neue Wege in der Sozialisationsforschung*, Kapitel 23. Stuttgart: Frommann-Holzboog.
- DREESMANN, H. (1982). *Unterrichtsklima - Wie Schüler den Unterricht wahrnehmen*. Weinheim und Basel: Beltz.
- DUTKA, W. (1985). *Die Wahrnehmung des Lehrerverhaltens durch die Schüler: Erforschung und Beeinflussung ausgewählter Unterrichtsklimamerkmale in der Grundschule - Bericht über ein Forschungsprojekt und dessen Ergebnisse*. Dissertation am Institut für Psychologie im Fachbereich 12 der Freien Universität Berlin.
- KAHL, TH., BUCHMANN M. & WITTE, E. (1977). Ein Fragebogen zur Schülerwahrnehmung unterrichtlicher Lernsituationen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9, 277-285.
- LEWIN, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.
- MURRAY, M.A. (1938). *Explorations in personality*. New York: Oxford University Press.
- MASENDORF, F., TÜCKE, M., KRETSCHMANN, R. & BARTRAM, M. (1976). *Dortmunder Skala zur Erfassung von Lehrerverhalten durch Schüler*. Braunschweig: Westermann.
- MOOS, R.H. (1974). *The social climate scales - an overview*. Palo Alto: Consultative Psychologists.