INVESTIGATION TO THE EFFECTS OF TWO EXTRA LESSONS IN PHYSICAL EDUCATION A WEEK DURING A SCHOOLYEAR UPON THE PHYSICAL AND MENTAL DEVELOPMENT OF 12 AND 13-YEAR OLD BOYS, PUPILS IN THE FIRST FORMS OF A SECONDARY SCHOOL IN AMSTERDAM

University of Amsterdam
Jan Swammerdam Institute
Eerste Constantijn Huygenstraat 20
Amsterdam - The Netherlands

Han C.G. Kemper
Koos G.A. Ras
Jan Snel
Paul G. Splinter
Louis W.C. Tavecchio
Robbert Verschuur

Coronel Laboratory
Director: Prof. Dr. R.L. Zielhuis
Laboratory of Psychophysiology
Director: Prof. Dr. P. Visser

For this project a grant had been awarded of the Foundation for Educational Research (S.V.O.) and the Ministry of Health and Environmental hygiene in The Hague, The Netherlands
(project-number 0185)

SUMMARY

Chapter 1, Introduction

Many experiments in the field of physical education have been performed to investigate the effects of extra lessons of physical education (scheme 1.1.). Reviewing this literature it can be stated that there was no systematic replication In order to retest results of a pilot study (Oud e.a., 1971) we decided to replicate this former experiment.

Chapter 2, Experimental design

The subjects were boys of the four first forms of a secondary school in Amsterdam, St. Ignatius College. From 82 boys, 12 boys dropped out (table 2.1.). At the beginning of the schoolyear 1971/1972 their mean chronological age was 12.5 ($\pm$ 0.4). Each class had its own teacher of physical education, who at the same time was the mentor of that class.

Chapter 3, Independent variable

The independent variable was the frequency of lessons a week in physical education. Two classes were assigned by lot as experimental, and two classes as control group. The usual number of 3 lessons of physical education a week were given to the control group, and five lessons of physical education a week to the experimental group. The two extra lessons of physical education had been added to the time table in such a way that the experimental group received 34 instead of 32 lessons a week.

A starting point for this study was the aims formulated in the 'Proposal Curriculum Government Schools' (Voorstel leerplan Rijksscholen, 1968). This curriculum was evaluated by series of objective performance tests. The two extra lessons of physical education can be seen as quantitative extension of the normal curriculum of physical education.

Chapter 4, Dependent variables

At the beginning (pretest) and at the end of the schoolyear (posttest) groups of dependent variables were measured as stated below:

- anthropometric variables (table 9.5.), measuring body build and body composition;
- physiological variables (table 9.7.), measuring aerobic power (W_{170}, 12 min. run walk), neuromuscular characteristics (strength, speed, flexibility) and ventilation (FEV%);
- gymnological variables (table 8.2.), measuring progress in performances of physical education;
- psychological variables, measuring group coherence (SAGS) (table 9.10.), neuroticism (ABV) (table 9.12.) achievement motivation (PMT-K) (table 9.15.), attitudes (SSD) (table 9.16. and appendix IV), schoolattitude (table 9.17.), preference and ranking of physical education (table 9.18.).

Based upon a review of the literature expectations were formulated about the effect of the independent variable upon the dependent variables.

Chapter 5, Interfering variables

Although an increase from three to five lessons a week seemed sufficient to find effects of extra lessons of physical education, there are some interfering variables that could mask these possible effects. As interfering variables were considered:

1. Biological age, measured as skeletal age by X-ray photography of hand and wrist (fig. 5.2. and table 5.3.);
2. Habitual physical activity, measured by pedometers (photo 5.9.) and table 5.10.) and a questionnaire (table 5.11 and appendix V);
3. Teaching behaviour of the four teachers of physical education, judged by experts in professional (scheme 5.12.) and social-emotional aspects(scheme 5.15.) and by judgements of the pupils (scheme 5.21.).

Chapter 6, Statistical methods; Chapter 7, Hypotheses

The systematic differences between experimental and control groups were analysed by the following statistical preparing steps:

a. the dependent variables were divided into groups that were supposed to be more or less independent from each other;
b. the grouped pretest data were factoranalysed for both experimental and control group; further, the representative variables, found in both groups, were used to formulate the hypotheses;
c. in each hypothesis it was specified which interfering variable(s) and which interaction(s) besides the independent variable could contribute to the difference in score of the experimental and control group of this representative (table 7.7.).

The mean difference scores of experimental and control group were compared by covariance-analysis while making allowance for the influence of the interfering variables. The interfering variable, teaching behaviour, had not been measured as data from subjects themselves and could for that reason not be taken up in the covariance-analysis.

Chapter 8, Results

The twelve hypotheses were tested by covariance-analysis (table 8.1.). Only the explosive arm strength (handgrip) proved to increase significantly ($p \leq .01$) in the experimental group with regard to the control group as a result of the extra lessons in physical education.

Series of gymnological tests were used to evaluate the physical education in a narrower sense at four points of time. Analysis of these data resulted in a significantly higher score ($p \leq .009$) of the experimental group in respect to the control group (fig. 8.3.).

An unreduced multivariate-analysis was executed for those groups of variables in which the factor analysis indicated non-identical structure of experimental and control group. No significant differences could be found in attitudes towards the ten SSD concepts as result of the two extra lessons in physical education. This result meant that the effect of extra lessons upon the concepts 'hobby' and 'sport', found in the pilot study, could not be confirmed. In respect to the variables SAGS, PMT-K, preference and rankorder of physical education, this group of variables as a whole appeared to differ significantly between experimental and control group. On further analysis of the difference in scores, especially the accepted (O) and given (G) scores of the SAGS appeared to cause the difference. Perhaps these scores might be related with a class-effect.

Chapter 9, Discussion

Because of the fact that the class-effect and/of teacher-effect could not be taken directly into the covariance-analysis, the possible influence of this interfering variable was investigated in an explorative way (scheme 9.1.). For each of the twelve representatives a ratio was calculated. This ratio consisted of the effect of the independent variable plus the class and/or teacher-effects between experimental and control group on the one hand and the class and/or teacher-effects within the experimental and control group on the other hand. For some of the variables this ratio was

small and one may suppose that the difference between experimental and control group can be explained partly by class and/or teacher-effects. In the explosive arm strength this ratio was very high so that there was no reason to assume any class and/or teacher-effects; therefore the results of the covariance-analysis concerning this variable were consolidated. A teacher-effect was supported by results from the different judgements of teaching behaviour, particularly the teachers of the two experimental classes were opposite to each other in a number of aspects.

As to the results of covariance-analysis, there are, because of statistical considerations, reasons to reckon with the possibility that the two extra lessons of physical education, biological age and habitual physical activity had been of minor importance on a number of dependent variables. The reason might be found in:

- the relatively small increase of only two extra lessons of physical education a week;
- the choice of the measured dependent and interfering variables;
- the low reliability of measuring the interfering variables;
- a possible influence of other, unknown factors.

In conclusion it can be stated that in general the expectations about the effects of two extra lessons of physical education upon 12 and 13-year old boys could not be confirmed.

In a five versus a three lessons-a-week physical education programme a significant increase of the achievement in physical education was found in the experimental group. However, from the other twelve hypotheses a significant increase in only the explosive arm strength could be shown in the experimental group with regard to the control group.

It will be stressed that these results do not give any implication for the evaluation of the regular curriculum of physical education.

The importance of the regular curriculum of physical education may reveal in the obtained data of the relative preference for physical education among the twelve other subjects. At the onset of the schoolyear the mean rankorder of preference was second or third for the subject physical education. Moreover, this relatively high preference became somewhat higher during the schoolyear,although at the same time the mean school attitude decreased.

INHOUDSOPGAVE

(contents)

			blz.
		Voorwoord (preface)	1
Hoofdstuk 1.		Inleiding (introduction)	5
		Literatuur	19
Hoofdstuk 2.		Experimentele opzet (experimental design)	24
	2.1.	Proefpersonen	25
	2.2.	Proefklassen	27
	2.3.	Leraren lichamelijke oefening	27
	2.4.	Lesrooster	28
	2.5.	Tijdstip van meting	29
		Literatuur	31
Hoofdstuk 3.		Onafhankelijke variabele (independent variable)	32
		Literatuur	36
Hoofdstuk 4.		Afhankelijke variabelen (dependent variables)	37
	4.1.	Antropometrische en fysiologische variabelen	40
	4.1.1.	Antropometrische variabelen	40
	4.1.2.	Fysiologische variabelen	44
	4.1.2.1.	Het aerobe vermogen gemeten met behulp van een fietsergometerbelasting	47

			blz.
	4.1.2.2.	Het aerobe vermogen gemeten met behulp van eenvoudige prestatietests	51
	4.1.2.3.	Het meten van de eigenschappen van het neuro-musculaire systeem met behulp van eenvoudige prestatietests	55
	4.1.2.4.	Het meten van de ventilatieprestatie	59
	4.1.3.	Gymnologische variabelen	60
		Literatuur	64
	4.2.	Psychologische variabelen	69
	4.2.1.	S.A.G.S.	69
	4.2.2.	Neurotische labiliteit	73
	4.2.3.	Bereidheid tot het leveren van prestaties	75
	4.2.4.	Attituden ten aanzien van concepten waarop de lichamelijke oefening invloed zou kunnen hebben	80
	4.2.5.	Schoolattitude	86
	4.2.6.	Gymnastiek, voorkeur en rangnummer	88
		Literatuur	93
Hoofdstuk 5.		Storende variabelen (interfering variables)	100
	5.1.	Biologische leeftijd	100
		Literatuur	112
	5.2.	Buitenschoolse lichamelijke activiteit	114
		Literatuur	122
	5.3.	Lesgeefgedrag van de leraren lichamelijke oefening	124
	5.3.1.	Vaktechnische beoordeling van het lesgeef-gedrag	126
	5.3.2.	Sociaal-emotionele beoordeling van het lesgeefgedrag	132
	5.3.3.	Activiteit tijdens de lessen lichamelijke oefening	153
	5.3.4.	Leerlingbeoordeling van het leraargedrag	153
		Literatuur	159
Hoofdstuk 6.		Statistische methoden (statistics)	162
	6.1.	Betrouwbaarheid	165
	6.2.	Factoranalyse	178
	6.3.	Covariantieanalyse	182
	6.4.	Niet gereduceerd twee-steekproeven-probleem	185
	6.5.	Vergelijking van de uitgevoerde analyse met een andere methode	186

	6.6.	Analyse van gymnologische variabelen	187
		Literatuur	188
Hoofdstuk 7.		Hypothesevorming (hypotheses)	191
	7.1.	Voorlopige hypothesen	193
	7.2.	Resultaten factoranalyse	195
	7.3.	Hypothesen getoetst door middel van covariantie	202
	7.4.	Hypothesen getoetst door middel van een niet gereduceerde multivariate analyse	204
	7.5.	Hypothese met betrekking tot de gymnologische variabelen	207
		Literatuur	208
Hoofdstuk 8.		Resultaten (results)	209
	8.1.	Resultaten covariantieanalyse	210
	8.2.	Resultaten niet gereduceerde multivariate analyse	212
	8.3.	Resultaten gymnologische variabelen	213
Hoofdstuk 9.		Discussie (discussion)	216
	9.1.	Covariantieanalyse	218
	9.2.	Niet gereduceerde multivariate analyse	223
	9.3.	Analyse gymnologische variabelen	225
	9.4.	Onafhankelijke variabele	227
	9.5.	Antropometrische en fysiologische variabelen	229
	9.6.	Psychologische variabelen	235
	9.7.	Biologische leeftijd	247
	9.8.	Buitenschoolse lichamelijke activiteit	248
	9.9.	Beoordeling lesgeefgedrag	252
		Literatuur	261
Samenvatting en conclusies (summary and conclusions)			263
Bijlagen (appendices):			270
	I	Brief van het Ministerie van onderwijs en wetenschappen	271
	II	Jaarplan lichamelijke oefening op de brugklassen van het St. Ignatius College	272
	III	Eenvoudige prestatietests	280
	VI	Standaard semantische differentiaal	294
	V	Vragenformulier buitenschoolse activiteiten en rookgewoonten	297

		blz.
	Hoofdstuk 1.	
Schema	1.1.	10
	Hoofdstuk 2.	
Tabel	2.1.	26
Tabel	2.2.	26
Tabel	2.3.	28
Schema	2.4.	28
Schema	2.5.	30
	Hoofdstuk 3.	
Tabel	3.1.	34
	Hoofdstuk 4.	
Schema	4.1.	48
Tabel	4.2.	53
Tabel	4.3.	54
Tabel	4.4.	58
Tabel	4.5.	63
Tabel	4.6.	84
Schema	4.7.	91
	Hoofdstuk 5	
Figuur	5.1.	101
Figuur	5.2.	104
Tabel	5.3.	107
Figuur	5.4.	108
Tabel	5.5.	109
Tabel	5.6.	110
Figuur	5.7.	110
Figuur	5.8.	111
Foto	5.9.	117
Tabel	5.10.	120
Tabel	5.11.	121
Schema	5.12.	128
Tabel	5.13.	130
Schema	5.14.	136
Schema	5.15.	139
Figuur	5.16.	144
Schema	5.17.	145
Tabel	5.18.	148

Tabel	5.19.	150
Tabel	5.20.	151
Tabel	5.21.	153
Schema	5.22.	156
Tabel	5.23.	157
	Hoofdstuk 6	
Tabel	6.1.	167
Tabel	6.2.	169
Tabel	6.3.	170
Tabel	6.4.	172
Schema	6.5.	173
Tabel	6.6.	175
	Hoofdstuk 7	
Tabel	7.1.	196
Tabel	7.2.	197
Tabel	7.3.	198
Tabel	7.4.	200
Tabel	7.5.	200
Tabel	7.6.	201
Schema	7.7.	205
	Hoofdstuk 8	
Tabel	8.1.	211
Tabel	8.2.	214
Tabel	8.3.	215
	Hoofdstuk 9	
Schema	9.1.	219
Tabel	9.2.	226
Tabel	9.3.	227
Tabel	9.4.	227
Tabel	9.5.	229
Tabel	9.6.	230
Tabel	9.7.	232
Tabel	9.8.	233
Tabel	9.9.	235
Tabel	9.10.	236
Tabel	9.11.	237
Tabel	9.12.	238

Tabel 9.13. 240
Tabel 9.14. 240
Tabel 9.15. 241
Tabel 9.16. 244
Tabel 9.17. 245
Tabel 9.18. 246
Tabel 9.19. 250
Tabel 9.20. 251
Tabel 9.21. 259

VOORWOORD

Bij het gereedkomen van dit rapport past een terugblik over de periode waarin dit onderzoek werd uitgevoerd. Hiermee kan tevens recht gedaan worden aan allen die op een of andere wijze een rol hebben gespeeld bij de realisering van dit project.

De voorbereiding startte reeds in augustus 1971 na beëindiging van het vooronderzoek dat door het Laboratorium voor Psychofysiologie en het Coronel Laboratorium van de Universiteit van Amsterdam op het St. Ignatius College werd uitgevoerd.

Aan het nieuwe onderzoeksplan werd door beide hoogleraren, Prof. Dr. P. Visser en Prof. Dr. R.L. Zielhuis, goedkeuring gehecht. Tijdens de gehele uitvoering van het onderzoek mochten wij onder hun supervisie alle mogelijke steun ondervinden. Van beide vakgroepen, zowel als van de werkgroep fysiologie en pathofysiologie van het menselijk bewegen, werd tijdens werkbesprekingen geprofiteerd van kritische opmerkingen en suggesties, de opzet en uitvoering betreffende.

Het onderzoeksplan werd ter subsidiëring aangeboden aan de Stichting Voor Onderzoek Van Het Onderwijs (S.V.O.). Door bemiddeling van

Drs. J. Haantjes, wetenschappelijk hoofdmedewerker aan de afdeling projectenbeheer van deze stichting, werd in september 1971 de goedkeuring voor de aanvullende subsidie verkregen.

Door het Ministerie van Sociale zaken en Volksgezondheid werd voor 50% in het te subsidiëren bedrag geparticipeerd.

Het bestuur van de St. Willibrordus Stichting richtte bij monde van haar voorzitter Drs. B.Bot s.j. op 8-8-1971 een verzoek aan het Ministerie van Onderwijs en Wetenschappen om in het schooljaar 1971/1972 bij twee brugklassen twee lessen lichamelijke oefening per week aan het normale lesrooster te mogen toevoegen.

Van het Ministerie van Onderwijs en Wetenschappen werd op 24-4-1972 toestemming ontvangen om van de lessentabel in het brugjaar af te wijken (zie bijlage I).

Onze bijzondere dank gaat uit naar de directie van het St. Ignatius College, met wier hulp de extra lessen in het rooster van de normale lessen opgenomen konden worden. De directie stelde voor ons onderzoek een laboratorium op school ter beschikking en ze nam verder alle ongerief dat door de experimentele opzet werd veroorzaakt, voor lief.

Woorden schieten tekort om te beschrijven wat de leraren lichamelijke oefening voor het onderzoek hebben betekend:

- gedurende een heel jaar van voorbereiding werd door hen een uniform leerplan voor het brugjaar opgesteld, en de inhoud van alle lessen uitgewerkt,
- tijdens de uitvoering van het experiment onderwierpen zij zich bij de uitvoering van het leerplan aan de zeer rigoreuze eisen om alle lessen aan de brugklassen zo uniform mogelijk te geven,
- bovendien stemden zij ermee in zich het hele schooljaar te laten bespieden met camera en microfoon, teneinde een beoordeling van hun lesgeefgedrag mogelijk te maken.

Aan Jos, Loek en Ed spreken wij daarvoor heel bijzonder onze waardering uit.

De heer E. Kupers, inspecteur lichamelijke opvoeding van de gemeente Amsterdam, verleende ons medewerking bij de keuze van inspecteurs voor de lichamelijke oefening die een aantal lessen, tijdens het onderzoek aan de brugklassen door de leraren gegeven, op hun vaktechnische aspecten beoordeelden. Mevrouw M.F. van der Putten en de heren J.M.

Berrevoets, P.C. Duinker, K. Postma, J. Kan, bleken bereid om een aantal middagen vele uren video-band aan zich te laten voorbijtrekken. Van Drs. H.D. Scholten kregen wij deskundige hulp bij het opstellen van een methode ter beoordeling van sociaal-emotionele aspecten van het lesgeefgedrag. Vele maandagavonden leidde hij de training van vijf beoordelaars: N. van Bergen, F. Meyers, P. Splinter, H. van de Weg en A. te Winkel die, naast hun studie, belangeloos zeer veel tijd en energie aan deze training besteedden en vervolgens van een groot uithoudingsvermogen blijkgaven bij de beoordeling van de ruim veertig lessen.

Het Mathematisch Centrum heeft een zeer belangrijke rol gespeeld door zijn adviezen over en uitvoering van de statistische bewerking van de resultaten. Onder leiding van Drs. R. Helmers werd door Drs. H. Elffers vanuit een inspirerende interesse met grote werklust en in intensief overleg met de onderzoekers naar optimale oplossingen gezocht. In samenwerking met hen werd ook een gedeelte van hoofdstuk 6 van dit rapport geschreven.

Voorts maakte Drs. J.J.L. Pieters, arts, alle röntgenfoto's van het handskelet met apparatuur van het Centraal Instituut voor Voedingsonderzoek (C.I.V.O., T.N.O.) te Zeist. Na vooraf door Dr. R. Steendijk te zijn ingewijd in de geheimen van beoordeling van röntgenfoto's, heeft mevrouw C. IJkel-Van Anraad de rond 200 foto's gescoord; mevrouw G. de Graaff, mevrouw L. de Vries en de heren J. Pijnenburg en W. Terpoorten verleenden op vele momenten de nodige technische en administratieve hulp, terwijl de heer K. Ras alle tests in zijn vrije tijd afnam. De heren K. Louwes en N. Brandjes verzorgden het fotomateriaal. Het definitieve manuscript werd 'persklaar' gemaakt door mevrouw G.A. de By-Klitsie.

Allen lijken kleine radertjes in het totale project, maar zonder hen zou dit project niet zo voorspoedig zijn verlopen.

Tot slot moeten wij een zeventigtal andere medewerkers bedanken voor de medewerking die zij verleenden, zonder te weten waarvoor zij die verleenden. Alle jongens uit de vier brugklassen van het schooljaar 1971/1972 hebben alle metingen, tests, fietsproeven en interviews gelaten ondergaan. De jongens van de twee experimentele klassen 'werkten' bovendien nog twee lesuren per week 'over'. Bij het verschijnen

van de resultaten zullen de meesten van hen reeds in de derde klas zitten en veel van hetgeen hen overkwam zal uit hun herinnering zijn gewist; onze dank is echter niet minder groot.

Wij spreken de wens uit dat dit rapport gelezen, becritiseerd en bediscussieerd zal worden.

Dr. Han C.G. Kemper
projectleider

---

hoofdstuk 1

INLEIDING

- probleemstelling
- eigen vooronderzoek
- ander onderzoek
- de onafhankelijke variabelen
- de afhankelijke variabelen
- de storende variabelen
- replicatie van het eigen vooronderzoek

LITERATUUR

1. INLEIDING

Probleemstelling

Het belang van lichamelijke activiteit voor het menselijk organisme in het algemeen en voor jeugdigen in het bijzonder, is algemeen aanvaard (Cureton, 1964; Parizkova, 1968; Astrand e.a., 1971; Shephard, 1973). Het hedendaagse leefpatroon kenmerkt zich echter door een grote bewegingsarmoede en een calorie-rijke voeding; bij schoolgaande kinderen wordt de bewegingsarmoede nog geaccentueerd door het langdurig moeten zitten in de klas. Om die redenen lijkt het noodzakelijk het activiteitspatroon van scholieren te verhogen zowel binnen schoolverband door middel van lessen lichamelijke oefening, als daarbuiten door middel van sport en andere vormen van actieve recreatie.

In alle geledingen van het onderwijs is het vak lichamelijke oefening in het rooster van lessen opgenomen. Dit geldt voor het basisonderwijs alsook voor elke vorm van voortgezet onderwijs. In Nederland varieert het aantal lessen dat wekelijks voor lichamelijke oefening in het lesrooster is opgenomen, van één tot drie lesuren van elk 45 tot 50 minuten. In het buitenland ligt dit aantal gewoonlijk niet veel anders.

De lichamelijke vorming neemt daarbij, gemeten aan het aantal lessen per week, een zeer bescheiden plaats in ten opzichte van de intellectuele vorming: een verhouding van 1 : 10. Men kan zich afvragen of dit aantal lessen en de wijze waarop deze lessen worden gegeven wel voldoende zijn indien niet alleen de intellectuele vorming maar ook de gehele persoon naar lichaam en geest ontwikkeld dient te worden.

Daarnaast kan men een opvatting signaleren dat lichamelijke fitheid samenhangt met geestelijke fitheid (mens sana in corpore sano). Psychosomatisch onderzoek heeft aangetoond dat lichamelijke veranderingen veroorzaakt kunnen worden door invloeden van psychische aard (Kimball, 1970). Vanuit de sport en de lichamelijke oefening lijkt het dan ook logisch te veronderstellen dat lichamelijke fitheid op haar beurt ook psychische veranderingen zal kunnen veroorzaken. Hoewel men algemeen aanneemt dat dit zo is, blijken ten aanzien van dit psychofysiologische aspect van lichamelijke activiteiten bijzonder weinig betrouwbare gegevens te zijn (Hammett, 1967). Enerzijds zijn

er wel dikwijls positieve correlaties aangetoond tussen het lichamelijke prestatievermogen en psychologische factoren, zoals prestatiemotivatie, attitudes en persoonlijkheidskenmerken, anderzijds echter kunnen deze correlaties ook veroorzaakt zijn door aanleg of voorkeur. In een literatuuroverzicht wijst Korting (1972) erop dat correlatie-onderzoekingen over de aard van de samenhang tussen motoriek en schoolprestaties op geen enkele wijze iets duidelijk maken.

Eigen vooronderzoek

In een vooronderzoek (Oud e.a., 1971) is getracht de invloed te bestuderen van twee extra lessen lichamelijke oefening per week op de lichamelijke en geestelijke ontwikkeling van 12 en 13 jarige jongens, brugklasleerlingen op een school voor VWO en HAVO te Amsterdam.

Het uitgangspunt van dit vooronderzoek vormde een aantal veronderstellingen en vermoedens met betrekking tot de gunstige invloed van lichamelijke oefening op zowel de lichamelijke als geestelijke groei en ontwikkeling van scholieren. Een aantal daarvan zijn geformuleerd in het Voorstel leerplan Rijksscholen voor lichamelijke oefening voor V.W.O., H.A.V.O. en M.A.V.O. (1968). De doelstelling van de lichamelijke oefening wordt daarin als volgt omschreven:

"Het onderwijs in de lichamelijke opvoeding heeft ten doel bij te
"dragen tot de vorming van de persoonlijkheid en is, met voortduren-
"de inachtneming daarvan, mede gericht op:
" . gunstige beïnvloeding van de lichamelijke ontwikkeling,
" . bevordering van een goede houdings- en bewegingsvorming,
" . vergroting van de bereidheid en geschiktheid tot het leveren
" van prestaties,
" . bevordering van de zin tot samenwerking,
" . hygiënische gewoontevorming,
" . kennismaking, mede met het oog op de besteding van de vrije
" tijd, met waardevolle vormen van actieve recreatie."

De resultaten van dit vooronderzoek kunnen als volgt worden samengevat: als gevolg van twee extra lessen lichamelijke oefening per week gedurende een heel schooljaar konden geen verbetering in anthropometrische en fysiologische kenmerken worden aangetoond. Ook het

aantal absenties blijkt niet te zijn verminderd. Van de schoolprestaties bleken de leerprestaties niet, maar de gymnastiekprestaties wel te zijn verbeterd.

Geen enkele invloed kon worden aangetoond op de persoonlijkheidsvorming. Echter een verandering van de attitude, in de zin van een grotere waardering van de actieve recreatie in de vrije tijd, kon duidelijk worden aangetoond.

Ander onderzoek

Het effect van lessen lichamelijke oefening op de lichamelijke en geestelijke groei en ontwikkeling van de schoolgaande jeugd is reeds vele jaren onderwerp van onderzoek. Omdat deze experimenten over het algemeen te onvolledig, te fragmentarisch en dikwijls ook te verschillend zijn van opzet, blijkt er nauwelijks sprake te zijn van enige overeenstemming in resultaten.

Het gebied van de lichamelijke oefening is van een zo ingewikkelde structuur dat het niet zonder meer toegankelijk is voor empirische methoden (Korting, 1972) en de resultaten van dergelijke experimenten zijn afhankelijk van een groot aantal factoren. Bij de bespreking van de diverse onderzoeken zullen wij achtereenvolgens aandacht besteden aan de volgende factoren:

1. factoren die samenhangen met de aard van het programma voor de lichamelijke oefening: de onafhankelijke variabele(n) in deze experimenten,
2. factoren die samenhangen met de aard van de kenmerken die men wil beïnvloeden: de afhankelijke variabelen,
3. factoren die samenhangen met de proefgroep die men onderzoekt: leeftijd, geslacht, sociaal-economisch milieu, intelligentieniveau, trainingstoestand, e.d.: deze factoren worden de storende variabelen genoemd omdat ze ertoe kunnen leiden dat een eventueel effect van de onafhankelijke variabele(n) op de afhankelijke variabele(n) niet manifest wordt.

De onafhankelijke variabelen

De onafhankelijke variabelen worden gevormd door het programma van de lichamelijke oefening. Wij kunnen daarbij onderscheidingen maken naar inhoud, intensiteit, duur en frequentie van de lessen per week.

Over het effect van lessen lichamelijke oefening met een verschillende inhoud is onderzoek gedaan door Rath (1942), Fabricius (1964), Ismail e.a. (1967) en Hladil (1966).

Het effect van lessen lichamelijke oefening met gelijke inhoud maar met verschillen in intensiteit is onderzocht door Dietrich (1966), Scheibe e.a. (1966), Cermàk (1970) en Kemper e.a. (1971).

In Zwitserland is gedurende de jaren 1960-1965 met de 'Aargauer und Berner Schulversuche' (cit. Bovend'eerdt, 1971) getracht het effect na te gaan van verschillen in duur van de lessen lichamelijke oefening door een verdeling over de week van zes halve uren te vergelijken met drie lessen van elk één uur (zie schema l.l.).

Verreweg de meeste onderzoeken hebben zich beziggehouden met het effect van verschillen in frequentie van het wekelijkse aantal lessen lichamelijke oefening. In schema l.l. is een overzicht gegeven van deze experimenten. Wanneer men de resultaten vergelijkt, blijken op vele punten tegenstrijdigheden te bestaan.

Op de eerste plaats moet verondersteld worden dat de inhoud van de lessen lichamelijke oefening tussen de diverse experimenten sterk verschilt; bij de meeste verslagen ontbreekt echter een beschrijving van de inhoud der lessen. Van een vergelijking van resultaten alleen op het effect van een verschil in frequentie van de lessen zal daarom nooit sprake kunnen zijn.

Op de tweede plaats blijken bij verschillende experimenten niet alleen de frequentie van het aantal lessen gevarieerd te worden, maar tegelijkertijd ook inhoud en intensiteit van de lessen lichamelijke oefening (Schleusing, 1964; Sprynarova, 1973; Cumming e.a., 1969; Buchberger, 1971).

Op de derde plaats is het aantal extra lessen lichamelijke oefening per week niet gelijk. Het aantal extra lessen per week varieert van één lesuur (Bar-Or e.a., 1973; Weingarten e.a., 1971) tot negen en een half lesuur (Cumming e.a., 1969). Bij de meeste experimenten is slechts één variatie in de lesfrequentie aangebracht, maar bij

investigation			subjects		
author(s)	country	year	sex	age (pretest)	number
Wendler	G.D.R.	1936	fem.	9 - 10	± 60
Encausse	France	1957	male	10 - 11	± 60
Groll	Austria	1965	male	9 - 10	8 classes
Schleusing	D.D.R.	1964	fem. a. male	10 - 11	± 160
Tillman	U.S.A.	1966	male	12 - 18	50
Prasilova	C.S.S.R.	1966	?	6 - 7	?
Soukupova	C.S.S.R.	1966	?	?	?
Holatova	C.S.S.R.	1966	?	6 - 7	?
Bakonyi	Hungary	1969	fem. a. male	?	210 classes
Isaac e.a.	Belgium	1969	male	11 - 12	58
Cumming e.a.	Canada	1969	male	11 and 15	?
La von Johnson	U.S.A.	1969	fem. a. male	14	284 and 459
Saunders e.a.	U.S.A.	1969	fem. a. male	14 - 15	894
Gabler	G.D.R.	1970	?	10 - 12	3 classes
Buchberger	C.S.S.R.	1971	fem. a. male	11 - 12	115
Lamers e.a.	Netherl.	1971	male	11 - 14	106
Oud e.a.	Netherl.	1971	male	12 - 13	50
Paschen	G.D.R.	1971	fem. a. male	6 - 7	8 classes
Rieckert e.a.	G.D.R.	1972a	male	11 - 12	2 classes
Rieckert e.a.	G.D.R.	1972b	fem.	15 - 16	2 classes
Rieckert e.a.	G.D.R.	1972c	male	11 - 12	22
Van der Hoeven	Netherl.	1973	male	9 - 10	150
Bar-Or e.a.	Israel	1973	male	9 - 10	92
Weingarten e.a.	Israel	1973	male	9 - 10	92
Sprynarova	C.S.S.R.	1973	male	11	39

Schema 1.1. Schematic review concerning extra lessons in physical

experimental design		results				
frequency of lessons a week	total period	anthropom. variables	Physiol. var.	gymnol. var.	psychol. var.	school perf.
2 - 3 : 6	3 years	+		+		+
$17\frac{1}{2}$: $3\frac{1}{2}$	1 and 3 years	+	+		+	+
5 : 2	1 year	-	-	+	+	±
2x2 hrs.sport extra	4 years		+			
extr.phys.fitn.pr.	1 year		+		-	
6 extra lessons	5 years		+			
6 hours	?	-	-			
6 : 3	3 years		-		-	
4 : 2	?				+	
7 : 2	3 years	+	+		±	-
$9\frac{1}{2}$: 2 : 0	1 year		-			
5 : 2 - 3		±	±			
yes,no, partly	4 years		±			
5 : 2			-	+		-
9 : 2	2 - 3 years		±			
2 : 0	$\frac{1}{2}$ year		-			
5 : 3	1 year	-	-	+	±	-
5 lessons of 20 min. : 0	2 years			+		
6 : 2	3 years		-	+		-
6 : 2	3 years		+	+		-
6 : 2	$1\frac{1}{2}$ year		-			
2:4:6:8:10	1 year		±			
2 : 3 : 4	$\frac{1}{4}$ year		-			
2 : 3 : 4	$\frac{1}{4}$ year				-	
5 : 3 : 2	7 years	+	±			

investigations to the effects of education (and/or sport); + = positive effect
- = no effect

Bar-Or e.a. (1973), Weingarten e.a. (1973), Cumming e.a. (1969) en Van der Hoeven (1973) zijn meerdere variaties in de lesfrequentie toegepast.

Op de vierde plaats speelt niet alleen het absolute verschil in extra lessen een rol van betekenis, maar ook de verhouding van het totale aantal lessen die over de diverse proefgroepen is aangebracht. Bij een verhouding over experimentele en controle groep van 5 : 3 (La von Johnson, 1969; Oud e.a., 1971) of van 2 : 0 (Lamers e.a., 1971) of van 4 : 2 (Bakonyi, 1969) is wel het aantal extra lessen gelijk, maar de procentuele vermeerdering niet.

Wanneer deze verhouding in het aantal wekelijkse lessen niet verschilt, zijn de resultaten wellicht verder afhankelijk van de periode waarover deze extra lessen gegeven worden. Bar-Or e.a. (1973) en Weingarten e.a. (1973) bekeken het effect van twee extra lessen over een betrekkelijk korte periode van drie maanden. Prasilova (1966) gebruikte een periode van vijf jaar om het effect van zes extra lessen na te gaan. Sprynarova (1973) analyseerde haar resultaten zelfs in een onderzoek dat zich uitstrekte over zeven jaar. Van een aantal andere experimenten ontbreken in de verslaggeving de gegevens over de duur in zijn geheel (Soukupova, 1966; Bakonyi, 1969; La von Johnson, 1969; Gabler, 1970).

Een volgend belangrijk verschil tussen de opgesomde experimenten wordt gevormd door de wijze waarop de extra lessen op school tot stand zijn gekomen. Men kan daarbij drie typen van onderzoek onderscheiden:

1. Experimenten waarbij de extra lessen lichamelijke oefening wekelijks aan het normale lesrooster op school worden toegevoegd, zoals dat het geval is geweest bij Gabler (1979), Rieckert e.a. (1972a, 1972b, 1972c) en Oud e.a., (1971).
2. Experimenten waarbij het aantal wekelijkse lessen is verhoogd ten koste van andere schoolvakken, zoals bij 'demi-temps' stelsels van Encausse (1957), Isaac e.a. (1969), Groll (cit. Bovend'eerdt, 1971) en Van der Hoeven (1973).
3. Wendler (1936) deed een experiment waarbij gedurende drie jaar de extra lessen slechts voor een deel ten koste van andere schoolvakken aan het rooster werden toegevoegd.

De evaluatie van de laatste twee typen onderzoek brengt een extra moeilijkheid met zich mee. Het is namelijk zeer moeilijk om ondubbelzinnig vast te stellen of de effecten veroorzaakt zijn door de vermeerdering van het aantal lessen lichamelijke oefening, of door een vermindering van de overige lessen. Korting (1972) stelt dan ook terecht dat in onderzoekingen zoals van Encausse (1957) en Isaac e.a. (1969) de door hen gevonden positieve resultaten niet op de eerste plaats gezocht moeten worden in de lichamelijke oefening als zodanig, maar meer in de geheel veranderde leefsfeer van de school.

De afhankelijke variabelen

In schema 1.1. zijn de resultaten van de opgesomde experimenten zoveel mogelijk overzichtelijk gemaakt door een verdeling naar anthropometrische variabelen, fysiologische variabelen, gymnologische variabelen, psychologische variabelen en leerprestaties.

Uit het overzicht blijkt dat niet alle onderzoekers het effect van de extra lessen op dezelfde kenmerken hebben nagegaan. Vergelijking van de resultaten is alleen om die reden al dikwijls niet mogelijk. Maar ook in het geval dat wel dezelfde kenmerken zijn onderzocht, blijken daarvoor dikwijls verschillende methoden te worden toegepast. De resultaten van Bakonyi (1969), die zijn gegevens uitsluitend heeft verzameld uit antwoorden van leerlingen op vragenlijsten, moeten dan ook in dat licht bekeken worden.

Slechts een klein aantal auteurs heeft het effect van extra lessen lichamelijke oefening nagegaan op de leerprestaties (Wendler, 1936; Encausse, 1957; Isaac e.a., 1969; Gabler, 1970; Oud e.a., 1971; Rieckert, 1972a, 1972b). Ook hierbij is voorzichtigheid geboden wanneer men hun resultaten wil vergelijken: Encausse (1957) neemt als criterium het aantal geslaagden voor een diploma, Groll de score op een intelligentietest (cit. Bovenä'eerdt, 1971) en de meeste anderen de schoolvorderingencijfers. De betrouwbaarheid van de meetmethode speelt derhalve een belangrijke rol bij het al dan niet vinden van positieve resultaten. Speciaal met betrekking tot de allereerste onderzoeken van Wendler (1936) en Encausse (1957) ontbreekt bovendien geheel een statistische toetsingsprocedure. De interpretatie van door

hen gevonden verschillen, als zijnde een positief resultaat, berust uitsluitend op gemiddelden, zonder rekening te houden met de spreiding van de gemeten kenmerken binnen de proefgroepen.

De storende variabelen

De resultaten van onderzoek naar het effect van extra lessen lichamelijke oefening worden niet uitsluitend bepaald door keuze van onafhankelijke en afhankelijke variabelen.

Allerlei persoonsvariabelen spelen eveneens een rol, zoals leeftijd en geslacht van de proefpersonen waarop men de onafhankelijke variabelen toepast. Maar ook de trainingstoestand en de interesse van de proefpersonen kunnen van betekenis zijn. De invloeden hiervan op de afhankelijke variabelen kunnen relatief zo groot zijn dat het effect van de onafhankelijke variabelen niet aantoonbaar wordt.

De evaluatie van training in het algemeen en lichamelijke oefening in het bijzonder is meestal afkomstig van transversaal onderzoek, waarbij een groep kinderen die geen lessen lichamelijke oefening krijgen, vergeleken wordt met een groep kinderen die wel een aantal lessen lichamelijke oefening krijgen. Door de keuze van de experimentele en controle groep kan men ernaar streven dat de diverse persoonsvariabelen gemiddeld niet verschillen (Schleusing, 1964; Saunders, 1969; La von Johnson, 1969). Het grote nadeel van een dergelijk 'static-group design' (Campbell e.a., 1966) blijft echter dat men de aanleg als eventueel storende variabele nooit goed in de hand heeft. Uit deze transversale experimenten is dan ook nooit met zekerheid vast te stellen of de gevonden verschillen berusten op verschillen in het aantal lessen lichamelijke oefening, of voor een deel veroorzaakt zijn door verschillen in aanleg.

Om die redenen hebben longitudinale onderzoekingen de voorkeur. Hierbij wordt het effect van de extra lessen nagegaan door de resultaten met de proefgroep aan het begin van de trainingsperiode te vergelijken met de resultaten aan het einde. Bij volwassenen, bij wie lichaamsbouw, lichaamssamenstelling en lichamelijke prestaties onder normale omstandigheden relatief constant blijven, mag men aannemen dat na een periode van extra lichamelijke activiteit en bij

overigens gelijkblijvende omstandigheden elke gevonden verandering daardoor is veroorzaakt. Bij kinderen echter levert dit "one-group-pretest-posttest design' weer andere problemen op: bij hen is de lichamelijke en geestelijke ontwikkeling aan voortdurende veranderingen onderhevig.

Voor wat betreft de lichamelijke aspecten werd door Ekblom (1969) en Von Döbeln e.a. (1973) een oplossing gevonden door in een longitudinaal onderzoek de gemeten veranderingen te vergelijken met veranderingen die men theoretisch uit de groei zou mogen voorspellen: wanneer de verschillen in de gemeten waarden in de loop van de trainingsperiode verder afwijken van de voorspelde waarden, mag aangenomen worden dat van een trainingseffect sprake is geweest.

Klissouras e.a. (1973) en Meller e.a. (1970) maakten gebruik van één-eiïge tweelingen, waarbij verschillen in aanleg als minimaal kunnen worden beschouwd.

Bij de meeste van de in schema 1.1. genoemde experimenten wordt gebruik gemaakt van longitudinaal onderzoek waarbij naast een experimentele groep tevens van een controle groep over dezelfde onderzoeksperiode wordt gebruik gemaakt.

Wanneer men, rekening houdend met leeftijd, geslacht en andere persoonsvariabelen, vergelijkbare proefgroepen heeft gekozen, mag men aannemen dat in dit 'pretest-posttest control group design' groei en ontwikkeling in beide groepen in dezelfde mate zullen plaatsvinden. Verschillen tussen beide groepen kunnen dan aan het einde van de trainingsperiode aan de extra lessen worden toegeschreven.

De wijze echter waarop bij vele van genoemde onderzoeken experimentele en controle groepen zijn samengesteld, vormt een belangrijk punt van critiek. Encausse (1957), Isaac e.a. (1969), Cumming e.a. (1969), Lamers e.a. (1971) en Paschen (1971) rekruteerden hun proefgroepen van verschillende scholen. Zij kunnen daarbij tussen experimentele en controle groepen verschillen introduceren in sociaal-economisch milieu en in andere persoonsvariabelen. Astrand e.a. (1971) wijst erop dat ook bij dergelijk longitudinaal onderzoek de resultaten vertekend kunnen worden door de volgende factoren:

1. Men maakt gebruik van vrijwilligers; bij Schleusing (1964), Buchberger (1971), Rieckert e.a. (1972a, 1972b, 1972c) en Sprynarova

(1973) is dat het geval. Het ligt voor de hand om te veronderstellen dat men, wanneer men vrijwilligers vraagt om extra lichamelijke activiteit te verrichten, door auto-selectie een groep kinderen krijgt met een grotere gemotiveerdheid en interesse en wellicht ook met een hoger lichamelijk prestatieniveau dan de gemiddelde populatie waaruit de groep afkomstig is.

2. Men werkt met groepen die reeds verschillen vertonen in trainingstoestand vóór de aanvang van het onderzoek; sommige auteurs introduceren dergelijke verschillen tussen hun proefgroepen door als experimentele groep leerlingen te kiezen met een laag lichamelijk prestatieniveau (Tillman, 1966) of met een hoog lichamelijk prestatieniveau (Gabler, 1970 , Rieckert e.a., 1972a, 1972b, 1972c).
3. Het percentage uitvallers tijdens een longitudinaal onderzoek: bij Sprynarova (1973) bleven van de 114 jongens van veertien jaar na vier jaar van longitudinaal onderzoek slechts 39 jongens over. Dit grote percentage uitvallers moet op zich zelf beschouwd worden als een min of meer geselecteerde groep (bijv. van kinderen die wat minder voor sport en lichamelijke activiteit gemotiveerd zijn dan de rest). Over de resterende 39 jongens valt derhalve geen algemene uitspraak meer te doen.

In het vooronderzoek (Oud e.a., 1971) wordt een aantal storende variabelen als oorzaak aangegeven voor het feit dat de twee extra lessen op de meeste afhankelijke variabelen geen effect hadden:

1. De proefpersonen hadden weliswaar ongeveer dezelfde kalenderleeftijd maar bij de bepaling van de ontwikkelingsleeftijd bleek de geslachtelijke rijping zeer ver uiteen te lopen. Dit impliceert tijdens het experiment een verschil in groei- en ontwikkelingssnelheid tussen de leerlingen.
2. De buitenschoolse lichamelijke activiteiten kunnen eveneens een storende variabele vormen. Men mag van twee extra lessen lichamelijke oefening bij zeer actieve jongens niet hetzelfde trainingseffect verwachten als bij zeer weinig actieve klasgenoten.
3. Voor motivationele factoren geldt hetzelfde: bij klassikaal onderwijs in de lichamelijke oefening zullen zij die deze lessen het meest appreciëren, het best opletten, het meest oefenen en daardoor de grootste vorderingen maken.

Verschillende auteurs wezen reeds eerder op de inter-individuele verschillen in ontwikkelingsleeftijd (Clarke, 1971) en dagelijkse lichamelijke activiteit (Rutenfranz, 1973). Ook Wendler (1936) en Van der Hoeven (1973) signaleren de verschillen in ectiviteitspatroon van hun proefpersonen, zonder echter in hun onderzoek daarmee rekening te houden.

Replicatie van het eigen vooronderzoek

Ondanks het feit dat de experimenten naar het effect van extra lessen lichamelijke oefening zeer talrijk zijn geweest, kan men toch stellen dat er nauwelijks sprake is geweest van een systematische herhaling van dergelijke experimenten. Een onderzoeker in de natuurwetenschappen zal eenzelfde experiment vele malen herhalen alvorens zijn bevindingen te publiceren. Op het gebied van de lichamelijke oefening ziet men nauwelijks enig replicatie-onderzoek (Pelton, 1970).

Omdat één enkel onderzoek zich praktisch niet leent tot generalisatie leek het ons gewenst om het eigen vooronderzoek te herhalen bij een vergelijkbare groep jongens van dezelfde school en met gebruikmaking van hetzelfde programma voor lichamelijke oefening.

Een nieuw element in dit vervolg-onderzoek vormt echter het feit dat de eerder gesignaleerde verschillen in ontwikkelingsleeftijd en dagelijkse lichamelijke ectiviteit tussen de leerlingen onderling niet alleen gemeten zullen worden maar tevens in een zodanig statistisch model zullen worden opgenomen dat daardoor het eigenlijke effect van de twee extra lessen lichamelijke oefening beter kan worden geschat.

Omdat het programma voor de lichamelijke oefening aan de vier brugklassen door vier verschillende leraren werd gegeven, bestaat de mogelijkheid dat een leraar-effect als storende variabele gaat optreden. Met o.a. video-opnamen zal daarom van een aantal van de gegeven lessen worden nagegaan in hoeverre de leraren ten aanzien van verschillende aspecten van hun lesgeefgedrag verschillen vertonen.

In het eigen vooronderzoek en in dat van Van der Hoeven (1973) werden de lessen aan alle klassen door dezelfde leraar gegeven om

zodoende een leraar-effect te vermijden. Wanneer echter in dit replicatie-onderzoek de verschillen in lesgeefgedrag niet al te groot zijn, kunnen de effecten van het programma voor de lichamelijke oefening dat in beide experimenten identiek was, enigszins gegeneraliseerd worden.

Literatuur

1. Astrand, P.O., K. Rohdahl, Textbook of work-physiology, New-York, McGraw-Hill, 1971.
2. Bakonyi, F., Der Einfluss der Schulischen Körpererziehung auf die Ansichten der Schüler über Körpererziehung und Sport, A testneveles tanitasa (Unterricht in Körpererziehung), Budapest, 1, 1969, 19-24.
3. Bar-Or, O., L.D. Zwiren, Physiological effects of increased frequency of physical education classes and of endurance conditioning on 9-10 year old girls and boys. In: Proc. 4th Int. Symp. on Ped. Work Physiol., Natanya, 1973, 183-198.
4. Bovend'eerdt, J., Onderzoek naar het effect van meer lich. opv. in de schoolprogramma's, Tilburg, intern rapport KALO, 1971.
5. Buchberger, J., Der Einfluss verschiedener Trainingsarten auf die Arbeitskapazität von Jugendlichen, Wien. Med.Wschr., 121, 51/52, 1971, 938-941.
6. Campbell, D.T., J.C. Stanley, Experimental and quasi-experimental designs for research, Rand Mc Nally and Comp., Chicago, 1966.
7. Cermák, J., Die Veränderungen des Herzvolumens und der somatometrischen Grundkriterien bei 12 bis 15 jährigen Knaben mit unterschiedlichem Bewegungsregime. XVIII World Congress of Sports Medicine, Oxford, 6-11 sept. 1970.
8. Clarke, H.H., Physical and motor tests in the Medford Boys' Growth Study, Prentice-Hall, 1971.
9. Cumming, G.R., D. Goulding, G. Baggley, Failure of school physical education to improve cardiorespiratory fitness, Canad. Med.Ass.J. 101, 2, 1969, 69-73.
10. Cureton, T.K., Improving the physical fitness of youth, Monographs of the Soc. Child Develop., 29, 4, 1964.
11. Dietrich, R., Die Beeinflussung des kardiorespiratorischen Systems bei Kindern durch unterschiedliche körperliche Belastung, Med. u Sport, 1966, 161-163.
12. Döbeln, W. von, B.O. Eriksson, Physical training, growth and maximal oxygen uptake of boys, aged 11-13 years, In: Ped. Work Physiology, Proc. 4th Int. Symp. on Work Physiol., Natanya, 1973, 93-108.
13. Ekblom, B., Effect of physical training on oxygen transport

system in man, Acta Phys. Scand., suppl. 328, 1969.

14. Eriksson, B.O., Physical training, oxygen supply and muscle metabolism in 11-13-year old boys, Acta Physiologica Scand., suppl. 384, Stockholm, 1974.
15. Encausse, P., Influence des activités physiques et sportives sur le développement intellectuel et physique en milieu scolaire, Paris, Imprimerie nationale, 1957.
16. Fabricius, H., Effect of added calisthenics on the physical fitness of 4th grade boys and girls, Res. Quart., 35, 1964, 135-140.
17. Gabler, H., Zur Einführung einer Sportzugklasse, Die Leibeserziehung, 1, 1970, 4-6.
18. Hammett, V.B.O., Psychological changes with physical fitness training, Canad. Med. Ass. J., 96, 1967, 764- 768.
19. Hladil, Z., Study of the influence of various contents of physical education and the level of the physical education process on the all around physical efficiency and functional fitness of apprentices, In: II Int. Congress about the physical fitness of youth, Praag, 1966, 435-440.
20. Holatová, D., Experiment to evaluate the influence of systematic physical education on the physical and mental state of youth, In: II, Int. Congress about the physical fitness of youth, Praag, 1966, 121-123.
21. Hoeven, H.C. van der, Lichamelijke oefening en bewegingsvaardigheid - een onderzoek naar de relatieve effectiviteit van een aantal in lesuren toenemende programma's lichamelijke oefening, scriptie voor de vakgroep ontw. psych., Univ. Nijmegen, 1973.
22. Isaac, J., C. Wilmart, F. van den Bossche, O. Marneffe, A. Colemont, H. Ide, E. van Kerchove, Het pedagogisch half-time stelsel van Cuesmes, Sport, 2, 46, 1969, 84-98.
23. Ismail, A.H., J.J. Gruber, Integrated development. Motor aptitude and intellectual performance, Columbus, Ohio, C.E. Merrill Books, 1967.
24. Johnson, C. la von, Effects of 5-day-a-week versus 2- and 3-day-a-week physical education class on fitness, skill, adipose tissue and growth, Res. Quart., 40, 1, 1969, 93-98.
25. Kemper, H.C.G., A.J. Poulus, N. van der Helm, Circuit-Training

und Körpererziehung, Ueber den Einfluss eines Kreistrainings beim Schulsport auf einige morphologische und funktionelle Merkmale bei 12- und 13-jährigen Jungen, Med. u Sport, 6, 1971, 179-184.

26. Kimball, C.P., Conceptual developments in psychosomatic medicine 1939 - 1969, Ann. intern. Med., 73, 1970, 307 - 316.
27. Klissouras, V., G. Weber, Training, growth and heredity. In: Proc. 4th. Inth. Symp. on Ped. Work Physiol., Natanya, 1973, 209-215.
28. Korting, F.W.J.P., De Lichamelijke Opvoeding, een bijdrage tot het bepalen van haar betekenis en plaats, scriptie Kath. Univ. Nijmegen, 1972.
29. Lamers-Winkelman, F., H.A. Keizer, Verkort verslag van een onderzoek naar de invloed van de lessen lichamelijke oefening op enkele fysiologische en somatometrische parameters, Thomas, 1, 1971, 2-10.
30. Meller, W., H. Mellerowicz, Vergleichende Untersuchungen über Dauertraining mit gleicher Arbeit, aber unterschiedlicher Leistung an eineiiger Zwillingen, Sportarzt und Sportmedizin, 1, 1970, 1
31. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor V.W.O. en H.A.V.O. op lichamelijke en geestelijke groei en ontwikkeling, werkstuk laboratorium voor psychofysiologie, Universiteit van Amsterdam, 1971, en in : Kemper, H.C.G. e.a., De Lich.Opv., 59, 12 en 13, 1971, 543 en 604.
32. Parizkova, J., Longitudinal study on the influence of physical activity on body build and the composition in boys from 11-15 years, Teor. Praxe tel Vykovy, Praag, 1968, 16, suppl. 31-34.
33. Paschen, K., Tägliche Bewegungszeit in der Grundschule; Bericht über das Forschungsvorhaben "Grundschulturnen" des Instituts für Leibesübungen der Universität Hamburg in den Jahren 67-69, Karl Hofmann, Schorndorff, 1971.
34. Pelton, B., A need for replication in research in health, physical education and recreation, Res. Quart., 41, 4, 1970, 613-615.
35. Prasilova, M., Influence of every day intensive physical education on children of the 1st - 5th grade, In: II Int. Congress about

the physical fitness of youth, Praag, 1966, 367-370.

36. Rath, E., A study of the effect of different physical education programs on the strength index of ninth grade boys, Res. Quart., 13, 1942, 169-177.
37. Rieckert, H., H. Gabler, H. Hinneberg, W. Schnezer, Drei Jahre tägliche Sportstunde in einer Mädchenklasse, Med. Welt, 23, 1972, 1115-1117 (a).
38. Rieckert, H., H. Gabler, F. Brandner, Vergleichende Untersuchungen im Schul- und Leistungssport, Arztliche Forschung, 36, 1972, 32-36 (b).
39. Rieckert, H., H. Gabler, Der Trainingseffekt einer täglichen Sportstunde auf das körperliche Leistungsvermögen von 11-12 jährigen Schulern, Sportarzt u. Sportmed., 2, 1972, 21-25 (c).
40. Rutenfranz, J., Physical performance capacity determined as W170 in Youth, In: Proc. 4th Int.Symp. on Ped.Work Physiol., Natanya, 1973, 245-249.
41. Saunders, R.J., H.J. Montoye, D.A. Cunningham, A.J. Kozar, Physical fitness of high school students and participation in physical education classes, Res. Quart., 40, 3, 1969, 552-560.
42. Scheibe, J., E. Blum, Der Einfluss intensiveren Turnunterrichts auf die körperliche Entwicklung von Schülern der 5. bis 7. Klassen, Med. u. Sport, 1966, 153-156.
43. Schleusing, G., H. Rieger, Längsschnittuntersuchung über die Wirkung des sportlichen Trainings auf den kindlichen Organismus vom 11. bis 14. Lebensjahr, Zeitschr. f. Alternforschung, 17, 3, 1964, 258-268.
44. Shephard, R.J., Challenges to fitness research, In: Proc. 4th Int. Symp. on Ped. Work Physiol., Natanya, 1973, 7-27.
45. Soukupova, D., Contributions to the observation of efficiency of students, In: II Int. Congress about the physical fitness of youth, Praag, 1966, 132-136.
46. Sprynarová, S., The impact of motor activity level on physical fitness during adolescence, Physical Fitness, Prague, 1973, 371-374.
47. Tillman, K., Relationship between physical fitness and selected personality traits, Res. Quart. 36, 1965, 483-489.
48. Weingarten, G., O. Bar-Or, Athletic and social status of elementary

school children - the effects of various contents of physical education programs, Thomas, 10, 1973, 329-343.

49. Voorstel leerplan Rijksscholen voor lichamelijke oefening voor V.W.O., H.A.V.O. en M.A.V.O., Ministerie van Onderwijs en Wetenschappen, Den Haag, 1968.

50. Wendler, F., Der dreijährige praktische Versuch mit der täglichen Turnstunde, Thesis, Hamburg, 1936.

hoofdstuk 2

EXPERIMENTELE OPZET

2.1. PROEFPERSONEN
- uitvallers
- leeftijd

2.2. PROEFKLASSEN

2.3. LERAREN LICHAMELIJKE OEFENING

2.4. LESROOSTER

2.5. TIJDSTIP VAN METING

LITERATUUR

2. EXPERIMENTELE OPZET

2.1. PROEFPERSONEN

De proefpersonen in dit onderzoek waren jongens uit de vier brugklassen van het St. Ignatius College, een school voor V.W.O. en H.A.V.O. te Amsterdam. In totaal werden in de brugklas 82 jongens en 20 meisjes aangenomen voor de cursus '71-'72. De klassen Gym. 1 en Gym. 2 werden gevormd door jongens en meisjes, aangemeld voor het Gymnasium; de klassen Ath/HAVO.3 en Ath/HAVO.4 door jongens en meisjes, aangemeld voor Atheneum en H.A.V.O..

Alle leerlingen voor de Gymnasiumafdeling werden aan het begin van het schooljaar verdeeld over beide parallelklassen Gym.1 en Gym.2, op basis van het alfabet. Dezelfde procedure werd gevolgd voor de indeling in de klassen Ath/HAVO.3 en Ath/HAVO.4 van de afdeling Atheneum/H.A.V.O.. Zittenblijvers werden gelijkelijk over de parallelklassen verdeeld.

De meisjes van de Gymnasiumklassen hadden, evenals de meisjes van de Atheneum/H.A.V.O. klassen, afzonderlijk gymnastiekles en werden daarom buiten het onderzoek gelaten. Desondanks werden ook zij getest wanneer het ging om klassikale test-situaties.

Uitvallers

Van de 82 jongens die bij de voortest deel uitmaakten van de brugklassen, bleven voor het onderzoek uiteindelijk 70 jongens over.

De criteria volgens welke een leerling als uitvaller werd beschouwd, waren:

1. afwezigheid op een der tijdstippen waarop de metingen plaatsvonden;
2. verzuim van meer dan 10% van alle lessen lichamelijke oefening die dat schooljaar werden gegeven.

Volgens deze criteria werden 12 jongens buiten het onderzoek gehouden, een uitvallerspercentage van 14,63%. Tabel 2.1. geeft een overzicht van de uitvallers per klas.

Composition of classes on pre- and posttest						
section	class	number of boys on pretest	drop-out frequency per criterium 1.	2.	total of drop-outs	number of boys on posttest
Gymnasium	Gym.1	19	0	3	3	16
Gymnasium	Gym.2	19	1	1	2	17
Atheneum/Havo	Ath/Havo.3	22	2	4	6	16
Atheneum/Havo	Ath/Havo.4	22	0	1	1	21
Gymnasium + Atheneum/Havo	1 + 2 + 3 + 4	82	3	9	12	70

Tabel 2.1. : Number of boys from the four first forms on pretest, drop-out frequency per criterium, and ultimate composition of the four first forms (posttest).

Leeftijd

De chronologische of kalenderleeftijd van de 70 jongens, die uiteindelijk in dit onderzoek overbleven, bedroeg aan het begin van het schooljaar gemiddeld 12,5 jaar (zie tabel 2.2.).

Voor gegevens met betrekking tot de biologische leeftijd van de proefpersonen, zie hoofdstuk 'storende variabelen' (5.1.).

Chronological age of the subjects (year decimals)					
	mean	s.d.	min.	max.	range
chronological age	12.5	0.4	11.9	13.7	1.8

Tabel 2.2. : Mean, standard deviation (s.d.), minimum, maximum and range of chronological age in year decimals of 70 boys from the four classes.

2.2. PROEFKLASSEN

De verdeling van de vier brugklassen over de experimentele en de controlegroep kwam tot stand door middel van loting. Om verschillen tussen de experimentele en de controlegroep zoveel mogelijk te beperken, werd op voorhand bepaald dat een klas bestaande uit leerlingen aangemeld voor het Gymnasium, evenals een klas bestaande uit leerlingen aangemeld voor de Atheneum/HAVO afdeling, tot de experimentele en de controlegroep dienden te behoren.

Uit voorgaand onderzoek (Kemper e.a., 1970; Oud e.a., 1971) is al gebleken dat het gebruik van twee parallelklassen op dezelfde school, zonder voorafgaande loting ter garantie van het verkrijgen van een a-selecte verdeling, gerechtvaardigd is. Aangezien tevens is gebleken dat de vergelijkbaarheid tussen Gymnasium en Atheneum/HAVO klassen minder goed is (Oud e.a., 1971), werden beide parallelklassen over de experimentele en de controlegroep verdeeld. De indeling in klassen als geheel werd om zijn vele praktische voordelen gehandhaafd.

De klassen Gym.2 en Ath/HAVO.3 vormden in dit onderzoek de experimentele groep; de andere klassen, Gym.1 en Ath/HAVO.4, vormden de controlegroep.

2.3. LERAREN LICHAMELIJKE OEFENING

De vier leraren lichamelijke oefening die ten tijde van het onderzoek aan het St. Ignatius College waren verbonden, kregen allen een brugklas toegewezen waaraan zij, als enigen, het gehele schooljaar lichamelijke oefening doceerden, terwijl zij tevens als klasseleraar optraden.

De klassen van de experimentele groep, Gym.2 en Ath/HAVO.3, kregen les van respectievelijk leraar B.en C.; de klassen van de controlegroep, Gym.1 en Ath/HAVO.4, van respectievelijk leraar A. en D. (tabel 2.3.).

Alle vier leraren waren in het bezit van de vereiste acte (M.O.p) en hadden meer dan vijf jaar lesgeefervaring. Ten tijde van het onderzoek bedroeg de gemiddelde leeftijd van de leraren dertig jaar (tabel 2.3.). Het leeftijdsverschil was klein, maximaal vier jaar.

De vier leraren waren reeds geruime tijd verbonden aan het St. Ignatius College en vormden een hechte vakgroep.

	age	class
teacher A.	28	Gym.1
teacher B.	30	Gym.2
teacher C.	32	Ath/HAVO.3
teacher D.	30	Ath/HAVO.4

Tabel 2.3. : Age of the four teachers in physical education, and their distribution over the four classes

2.4. LESROOSTER

De brugklassen hebben, dankzij de coördinatie van de leraren per vakgroep, hetzelfde lesrooster en per vak hetzelfde jaarplan. Alle leerlingen volgen dus het eerste jaar eenzelfde lesprogramma (schema 2.4.). Door deze werkwijze wordt het de leerling gemakkelijker gemaakt om aan het einde van het schooljaar over te stappen naar de Atheneum, H.A.V.O., of Gymnasium afdeling.

program of lessons in the first forms			
subject	number of lessons a week	subject	number of lessons a week
Dutch	4	drawing	2
French	4	physical education	3
English	3	music	1
Mathematics	4	manual labour	1
Biology	2	religion	1
Geography	2	study lessons	3
History	2		

Schema 2.4. : Number of lessons during the week per subject in the first forms

Het schooljaar 1971-1972 bestond uit 36 weken verdeeld in vier perioden: de periode tot de herfstvakantie, tot de kerstvakantie, tot de paasvakantie en tot de zomervakantie. Ten tijde van het onderzoek werden in de brugklassen 32 lesuren per week gegeven, waarbij elke les 50 minuten duurde. De leerlingen hadden school van maandag tot en met vrijdag. De lessen waren tussen 8.10 uur en 15.35 uur, met uitzondering van de woensdagmiddagen; dan was men vanaf 12.30 uur vrij.

Per week werden aan de controlegroep de normale drie lessen lichamelijke oefening gegeven. De experimentele groep kreeg totaal vijf lessen lichamelijke oefening per week. De twee extra lessen lichamelijke oefening werden aan het rooster toegevoegd zonder dat daarvoor andere lessen kwamen te vervallen. Daardoor kwam voor de experimentele groep het totaal aantal lessen per week op 34.

Voor de lessen lichamelijke oefening waren op school ter beschikking: twee modern ingerichte gymnastiekzalen en een tegelplateau, geschikt voor diverse zaalspelen en loopvormen uit de atletiek. Voor het zwemonderwijs waren twee zwembaden beschikbaar, het AMVJ- en het Zuiderbad, enkele minuten gaans van de school.

2.5. TIJDSTIP VAN METING

Bij de experimentele groep kan de invloed van de onafhankelijke variabele - in dit geval twee extra lessen lichamelijke oefening per week gedurende het schooljaar - op de afhankelijke variabelen, alleen worden vastgesteld door de gemeten verschillen te vergelijken met de verschillen die in dezelfde periode bij de controlegroep zijn opgetreden.

Dit betekent dat de metingen van de afhankelijke variabelen direct voor en na de experimentele periode dienen plaats te vinden.

De afhankelijke variabelen werden dan ook zowel direct aan het begin, als aan het einde van het schooljaar gemeten. De tijdstippen waarop de voor- en natest plaatsvonden, zijn vermeld in schema 2.5.. Tevens zijn hierin de meettijdstippen van de storende variabelen opgenomen.

De meting van de gymnologische variabelen vond plaats in de loop van het schooljaar.

testprogram dependent variables		
dependent variables	moments of time pretest	posttest
anthropometric var.	16 Aug. 1971	29 May - 7 Jun 1972
physiological var. simple performance tests	16 Aug. 1971	1 - 9 June 1972
physiological var. aerobic power with bicycle ergometer	between 1 Oct. and 1 Nov. 1971	between 24 April and 19 May 1972
psychological var. S.A.G.S.	24 Sept. 1971	1 and 2 June 1972
psychological var. A.B.V.	17 Aug. 1971	9 - 19 June 1972
psychological var. P.M.T.K.	17 Aug. 1971	9 - 19 June 1972
psychological var. S.S.D.	17 Aug. 1971	9 - 19 June 1972
psych. var. school attitude	24 Sept. 1971	1 and 2 June 1972
psych.var.gymnastics preference	24 Sept. 1971	1 and 2 June 1972

testprogram interfering variables		
interfering variables	moments of time first measurement	second measurement
biological age	16 Aug. 1971	30 May a. 6 June 1972
pedometers	between 1 Nov. and 1 Dec. 1971	between 6 and 29 March 1972
questionnaire activity	between 1 Oct. and 1 Nov. 1971	between 24 April and 19 May 1972

testprogram gymnological variables	
number of tests	period of testing
7	betw. beginning of schoolyear and autumn holidays
7	between autumn- and Christmas holidays
9	between Christmas- and Easter holidays
8	between Easter- and summer holidays

Schema 2.5. : Points of time on which dependent, interfering and gymnological variables have been measured.

Literatuur

1. Kemper, H.C.G., A.J. Poulus, N. van der Helm, Circuit-Training und Körpererziehung, Ueber den Einfluss eines Kreistrainings beim Schulsport auf einige morphologische und funktionelle Merkmale bei 12- und 13-jährigen Jungen, Med. u. Sport, 6, 1971, 179-184.

2. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor V.W.O. en H.A.V.O. op lichamelijke en geestelijke groei en ontwikkeling, werkstuk laboratorium voor psychofysiologie, Universiteit van Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv., 59, 12 en 13, 1971, 543 en 604.

hoofdstuk 3

ONAFHANKELIJKE VARIABELE

- jaarplan, leerplan, testplan
- lesplan

LITERATUUR

3. ONAFHANKELIJKE VARIABELE

Als onafhankelijke variabele is de frequentie van het aantal lessen lichamelijke oefening per week gehanteerd.

Gedurende het schooljaar 1971/1972 werden aan de controlegroep per week drie lessen lichamelijke oefening, en aan de experimentele groep vijf lessen lichamelijke oefening gegeven.

De vier brugklassen van het St. Ignatius College te Amsterdam, waaruit onze proefpersonen afkomstig waren, hebben normaal per week drie lessen lichamelijke oefening van elk vijftig minuten in hun lesrooster.

Jaarplan, leerplan, testplan

Aan de controlegroep werden de drie lessen lichamelijke oefening gegeven op drie verschillende dagen verdeeld over de week. De inhoud van deze lessen was volgens een jaarplan, in voorafgaande jaren opgesteld door de vakgroep lichamelijke oefening, bestaande uit vier leraren lichamelijke oefening van deze school.

Als richtlijn voor dit jaarplan diende het Voorstel leerplan Rijksscholen (1968), waarin voor de brugklas naast gymnastiek-, spel- en atletieklessen ook zwemlessen worden geadviseerd. Het jaarplan bestaat uit een leerplan en een testplan.

In het leerplan is telkens voor een periode van zes weken een aantal eindvormen vastgelegd. Deze eindvormen fungeren als eindstations voor bepaalde bewegingsvormen; de leraren kunnen deze echter bereiken met verschillende methodieken.

In het testplan zijn in elke periode voor een aantal van de eindvormen objectieve prestatietests opgesteld om het leerresultaat te kunnen meten (zie 4.1.3. gymnologische variabelen).

Twee van de drie lessen per week werden volgens dit jaarplan besteed aan gymnastiek-, spel- en atletiekonderwijs en één les per week aan zwemonderwijs.

Omdat het zwemmen alleen gerealiseerd kon worden in het schooljaar waarin het experiment plaatsvond, zijn in het jaarplan hiervoor geen afzonderlijke eindvormen opgenomen.

De zwemlessen van de vier brugklassen vonden tegelijk plaats; er werd geoefend in vier groepen waarbij het klasseverband werd verlaten om gedifferentieerd naar vaardigheid te kunnen lesgeven.

Het leerplan voor zwemmen was geheel volgens het Voorstel leerplan Rijksscholen (1968) en het testplan bestond uit twee zwemtests die herhaalde malen tijdens het schooljaar werden afgenomen.

Het cijfer voor lichamelijke oefening werd bepaald aan de hand van de resultaten van alle afgelegde gymnologische en zwemtests.

De twee extra lessen lichamelijke oefening voor de experimentele groep vormden uitsluitend een uitbreiding van het totale aantal lessen lichamelijke oefening in kwantitatieve zin; deze twee lessen werden derhalve, over het hele schooljaar gerekend, als volgt verdeeld:

- $1\frac{2}{3}$ extra les werd besteed aan gymnastiek, spel en atletiek;
- $\frac{1}{3}$ extra les werd besteed aan zwemmen.

Op deze wijze werd bereikt dat een evenredige vermeerdering van het aantal lessen voor de diverse vormen van onderwijs tot stand kwam (zie tabel 3.1.).

Aan de experimentele groep werden de vijf lessen lichamelijke oefening gegeven op de vijf schooldagen van de week, zodat er sprake was van één dagelijks lesuur lichamelijke oefening.

		p e r i o d t i l l				
	lessons of phys.educ.	autumn holidays	X-mas holidays	Easter holidays	summer holidays	total number of lessons
normal number of lessons (3)	gymn/athl/ games	17	14	22	18	71
	swimming	8	7	11	8	34
extra number of lessons (2)	gymn/athl/ games	16	13	16	15	60
	swimming	-	2	8	-	10

Tabel 3.1. : Number of normal and extra lessons in physical education, during this investigation given to the experimental and control group, specified for the four periods and for lessons in gymnastics + athletics + games as well as in swimming.

Lesplan

Uitsluitend ten behoeve van het experiment werden door de vier leraren lichamelijke oefening in onderling overleg alle volgens het opgestelde jaarplan te geven lessen in het schooljaar voorafgaand aan het experiment (cursus 1970/1971), uitgeschreven en in een kaartsysteem bijeengebracht.

Bij de uitwerking van het lesplan is de volgende indeling gehanteerd:

1. inleiding (5 min.);
2. eerste deel (15 min.) met het accent op het aanleren van bewegingen;
3. tweede deel (15 - 20 min.) met het accent op doorgaande beoefening en prestatievorming, eventueel inclusief een slot.

Deze uitgewerkte lessen werden besproken en beproefd. Vervolgens werd de volgorde van deze lessen bepaald in het totale jaarplan.

Bovendien maakten alle tests van het testplan deel uit van het lesplan.

Aan het Ministerie van onderwijs en wetenschappen en aan de directie van het St. Ignatius College werd tevoren toestemming gevraagd om in twee van de vier brugklassen wekelijks twee extra lessen lichamelijke oefening aan het rooster te mogen toevoegen (bijlage I).

Welke twee klassen als experimentele en welke twee klassen als controlegroep zouden gaan fungeren, werd vooraf door het lot bepaald. Om echter te voorkomen dat de experimentele groep zich 'experimentele groep' zou gaan voelen, werden de volgende maatregelen getroffen:

1. ouders noch leerlingen werden op de hoogte gebracht van het experiment;
2. alle noodzakelijke metingen vonden plaats op school en werden verricht met hulp van de leraren lichamelijke oefening;
3. aan de controlegroep werd hetzelfde aantal metingen verricht als aan de experimentele groep;
4. de twee extra lessen lichamelijke oefening voor de experimentele groep waren aan het begin van het schooljaar opgenomen in het lesrooster, - omdat het hier alleen brugklasleerlingen betrof die voor het eerst naar het VWO en HAVO komen, zijn zij niet op de hoogte van het 'normale lesrooster en ervaren de extra lessen dus niet als 'extra'.

Literatuur

1. Voorstel leerplan Rijksscholen voor lichamelijke oefening voor V.W.O., H.A.V.O. en M.A.V.O., Ministerie van Onderwijs en Wetenschappen, Den Haag, 1968.

hoofdstuk 4

AFHANKELIJKE VARIABELEN

4.1. ANTROPOMETRISCHE EN FYSIOLOGISCHE VARIABELEN

4.1.1. ANTROPOMETRISCHE VARIABELEN

- invloed van beweging op bouw en samenstelling van het lichaam
- antropometrische beoordeling van bouw en samenstelling
- bepaling skeletbouw
- bepaling vetmassa
- bepaling spiermassa
- verwachtingen

4.1.2. FYSIOLOGISCHE VARIABELEN

4.1.2.1. HET AEROBE VERMOGEN GEMETEN MET BEHULP VAN EEN FIETSERGOMETERBELASTING

- keuze van het belastingsprogramma
- 'physical working capacity' (W 170)

- fietsergometer
- proefpersonen, registratie- en belastingsprocedure
- verwachting

4.1.2.2. HET AEROBE VERMOGEN GEMETEN MET BEHULP VAN EENVOUDIGE PRESTATIETESTS

- samenhang tussen veld- en laboratorium-tests
- testkeuze en meetprocedure
- verwachting

4.1.2.3. HET METEN VAN DE EIGENSCHAPPEN VAN HET NEURO-MUSCULAIRE SYSTEEM MET BEHULP VAN EENVOUDIGE PRESTATIETESTS

- testkeuze
- verwachtingen

4.1.2.4. HET METEN VAN DE VENTILATIEPRESTATIE

- verwachting

4.1.3. GYMNOLOGISCHE VARIABELEN

- vooronderzoek
- gymnologische tests
- verwachting

LITERATUUR

4.2. PSYCHOLOGISCHE VARIABELEN

4.2.1. S.A.G.S.

- de sociometrische methode
- keuze van het instrument
- verwachting

4.2.2. NEUROTISCHE LABILITEIT

. het instrument
. verwachtingen

4.2.3. BEREIDHEID TOT HET LEVEREN VAN PRESTATIES

. prestatiemotivatietest voor kinderen
. verwachtingen

4.2.4. ATTITUDEN TEN AANZIEN VAN CONCEPTEN WAAROP DE LICHAMELIJKE OEFENING INVLOED ZOU KUNNEN HEBBEN

. semantische differentiaal
. keuze van de concepten
. verwachtingen

4.2.5. SCHOOLATTITUDE

. ander onderzoek
. keuze van het instrument
. verwachting

4.2.6. GYMNASTIEK, VOORKEUR EN RANGNUMMER

. ander onderzoek
. vooronderzoek
. verwachting

LITERATUUR

4. AFHANKELIJKE VARIABELEN

4.1. ANTROPOMETRISCHE EN FYSIOLOGISCHE VARIABELEN

4.1.1. ANTROPOMETRISCHE VARIABELEN

De lichamelijke opvoeding stelt zich ten doel de lichamelijke ontwikkeling van de leerlingen gunstig te beïnvloeden. Om na te gaan in hoeverre die doelstellingen gerealiseerd worden wanneer zelfs twee lessen lichamelijke oefening aan het lesrooster worden toegevoegd, dient de lichaamsbouw meetbaar gemaakt te worden. Met behulp van antropometrische meetmethoden kan een indruk worden verkregen van de bijdrage van de diverse weefsels aan de lichaamsamenstelling.

Invloed van beweging op bouw en samenstelling van het lichaam

De bouw en samenstelling van het menselijk lichaam wordt voornamelijk bepaald door de hoeveelheid skelet-, spier- en vetweefsel. Alhoewel onder overigens 'normale' omstandigheden erfelijke factoren een overheersende rol spelen bij het bepalen van iemands algemene morfologische kenmerken, is bekend dat met name de ontwikkeling van spier- en vetweefsel wordt beïnvloed door uitwendige factoren zoals lichaamsbeweging en voeding. Aangezien wij in dit onderzoek te maken hebben met 12 en 13-jarigen, zal hun lichaamsbouw veranderingen ondergaan onder invloed van groeiprocessen. Het is dan ook niet uitgesloten dat de lichamelijke activiteiten zowel op de groei als op de ontwikkeling van de verschillende weefsels van invloed is.

Een minimum aan spierarbeid is een voorwaarde voor een normale groei van bot- en spierweefsel (Rarick, 1960; Malina, 1969). Hoe groot dat minimum moet zijn, is echter niet bekend. De bewegingsdrang van kinderen is over het algemeen echter dermate groot dat dit minimum ruimschoots zal worden overschreden.

Tevens is bekend dat vooral activiteiten met een hoge intensiteit een toename van de spierdoorsnede en de spierkracht veroorzaken (Rarick, 1960; Malina, 1969). Deze veranderingen zijn echter van tijdelijke aard en verdwijnen zodra men de training beëindigt. Gelijktijdig met deze toename in spierweefsel vindt een afname van de hoeveelheid vetweefsel plaats (Parizkova, 1963; Rarick, 1960; Malina, 1969). De oorzaak van deze verandering is gelegen in de verhouding energieopname en energieafgifte. Verhoogde lichamelijke activiteit veroorzaakt een

toename van de vetvrije massa ten koste van aanwezig vetweefsel. Het omgekeerde vindt plaats indien de mate van activiteit vermindert (Parizkova, 1963; Malina, 1969).

Ook bij onderzoek in Nederland naar de voedingstoestand van diverse bevolkingsgroepen constateerde men dat groepen met een grotere hoeveelheid dagelijkse lichamelijke activiteit minder lichaamsvet en meer spiermassa hebben dan de minder actieve groepen (Luyken, 1969 b en c). Bij een groep jongelui van 15 - 25 jaar werden deze veranderingen reeds waargenomen na afloop van een vierweekse cursus van zware lichamelijke activiteit (Luyken, 1969 a).

Op grond van bovenstaande gegevens kan in dit onderzoek verwacht worden dat bij meer lichamelijke oefening de vetmassa zal afnemen en de spiermassa zal toenemen.

Over de invloed van lichamelijke inspanning op het skelet is veel minder feitenmateriaal beschikbaar. Het ziet er echter naar uit dat lichamelijke inspanning tijdens de groeifase vooral de breedtegroei stimuleert (Rarick, 1960; Biener, 1961; Parizkova, 1963; Malina, 1969), terwijl in de lengtegroei weinig verandering optreedt (Rarick, 1960; Parizkova, 1963; Malina, 1969). Volgens Biener zal de lengtegroei door lichamelijke oefening alleen maar vertraagd worden, echter zonder dat de uiteindelijke lichaamslengte daardoor beïnvloed wordt. Dit effect treedt volgens Malina alleen op indien er grote druk op de botten wordt uitgeoefend. In de botstukken die niet betrokken zijn bij het dragen van het lichaam lijkt ook de lengtegroei door inspanning gestimuleerd te worden (Rarick, 1960; Malina, 1969).

Alhoewel de evidentie voor veranderingen in de groei van het skelet onder invloed van lichamelijke inspanning, met name de lengtegroei, niet groot is, verwachten wij toch een toename in de lengte zowel als in de breedtematen van het skelet, als gevolg van extra lessen lichamelijke oefening.

Antropometrische beoordeling van bouw en samenstelling

De basis voor een beoordeling van de lichaamsbouw wordt gevormd door het lichaamsgewicht en de lichaamslengte bij een gegeven leeftijd. Deze maten geven echter onvoldoende informatie.

Hoewel het lichaamsgewicht afhankelijk is van de verhouding waarin skelet-, spier- en vetweefsel voorkomen, geeft het geen informatie

over de afzonderlijke weefselmassa's. Vooral als men nagaat dat de variantie van het lichaamsgewicht van jonge volwassenen voor ongeveer 30% wordt verklaard uit de variabiliteit van spier- en vetmassa tezamen (Tanner, 1964), is duidelijk dat het lichaamsgewicht alléén een te grove maat vormt. Daar komt bij dat de vetmassa, maar vooral de spiermassa, op jongere leeftijd nog aan sterke veranderingen onderhevig is.

Bepaling skeletbouw

Als maat voor de bouw van het skelet geeft de lichaamslengte onvoldoende informatie. Factoranalytisch onderzoek heeft uitgewezen dat er behalve een factor skeletgrootte (lengte) een aantal factoren zijn die vormgeven aan de skeletbouw (Luyken, 1967). Terwijl de skeletgrootte wordt bepaald door lichaamslengte, romplengte of beenlengte, zijn het vooral de breedtematen zoals schouder-, borst- en bekkenbreedte, alsmede de botbreedte die de vorm aanbrengen. Vooral de breedte van de pijpbeenderen, gemeten op röntgenfoto's of als condylenwijdte, wordt beschouwd als de belangrijkste component van het skelet naast de skeletgrootte (Tanner, 1964). Deze breedtefactor verklaart ongeveer 10% van de variantie van alle lichaamsafmetingen, terwijl de skeletgrootte en botdikte tezamen 25-30% van de variantie voor hun rekening nemen (Tanner, 1964).

Ter benadering van de algehele groei en ontwikkeling werden in dit onderzoek zowel aan het begin als aan het einde van het schooljaar door getrainde personen de volgende metingen verricht:

- totaal lichaamsgewicht (kg) met behulp van een weegschaal van van Vucht's Concern, type Spido;
- lichaamslengte (mm), met behulp van een in het Laboratorium voor Psychofysiologie geconstrueerde digitaal afleesbare lengtemeter.

Ter benadering van de breedte van het skelet werden de volgende metingen verricht:

- schouder- en bekkenbreedte (mm), met behulp van een digitaal afleesbare meetlat, de Harpenden Antropometer;
- breedte van pols en knie (mm), met behulp van een Mauser schuifmaat, waarvan de bekranden waren verbreed.

Bepaling vetmassa

Voor het bepalen van de hoeveelheid vetmassa moet men zich het menselijk

lichaam voorstellen als de som van twee min of meer onafhankelijke compartimenten, namelijk de vetmassa en de vetvrije massa (v.v.m.). De vetmassa omvat al het lichaamsvet ongeacht de plaats en structuur. De vetvrije massa omvat de massa actief functionerend weefsel, te weten spier-, skeletmassa en de resterende weke delen, inclusief het bloed.

Om de hoeveelheid vetmassa of vetvrije massa te bepalen, kan gebruik gemaakt worden van een aantal methoden. De meest bekende methoden vormen de densitometrie en de antropometrie.

Bij de densitometrie berekent men de lichaamsdensiteit of -dichtheid door het lichaamsgewicht te delen door het volume. Het volume kan hierbij worden afgeleid door directe meting van de waterverplaatsing door het lichaam, of indirect via de bepeling van het lichaamsgewicht onder water. Aangezien de densiteit van spier-, skelet- en vetmassa verschilt, zal de totale lichaamsdichtheid uiteindelijk bepaald worden door de onderlinge verhouding van genoemde weefselsoorten. Via de meting van de densiteit kan men het aandeel vetmassa van het totale lichaamsgewicht berekenen (Brozek, 1963).

Aangezien de bepaling van de densiteit een gecompliceerde en kostbare methode vormt, heeft men getracht verbanden te vinden met een aantal antropometrische variabelen om zodoende de vetmassa of vetvrije massa met behulp van deze relatief eenvoudige metingen te kunnen schatten. Alhoewel men erin slaagde hoge correlaties te vinden tussen de densitometrie en de antropometrische voorspellingen van vetmassa en vetvrije massa, dient beseft te worden dat de hoge voorspellende waarde van de diverse formules alleen opgaat wanneer ze worden toegepast op eenzelfde populatie als waarop de formule oorspronkelijk was gebaseerd (Wilmore, 1969; Pieters, 1972). De onderwaterweegmethode of een vergelijkbare methode voor het meten van het lichaamsvolume met daarbij de directe berekening van het residuaire longvolume wordt algemeen als de meest nauwkeurige meting beschouwd.

In het vooronderzoek (Oud e.a., 1971) is bij 12 en 13-jarige jongens uit de brugklassen van het St. Ignatius College zowel de onderwaterweegmethode als de antropometrie toegepast. De regressievergelijking die uit deze gegevens werd opgesteld, maakte het mogelijk ditmaal te volstaan met antropometrische metingen. Uit twee huidplooien, triceps en subscapularis, kan het percentage vet worden geschat ($r = 0,87$) (Kemper, 1970, unpubl.).

In dit onderzoek werden ter benadering van de vetmassa de volgende metingen verricht:

. huidplooi van biceps, triceps, subscapularis en suprailiaca (mm), met behulp van een huidplooidiktemeter, de Harpenden Skinfold Caliper.

Bepaling spiermassa

De bijdrage van de spiermassa aan de lichaamssamenstelling kan eveneens langs indirecte weg met behulp van antropometrische metingen worden bepaald. Door gebruik te maken van omtreksmaten van armen of benen, als maat voor spier- en vetmassa tezamen, en huidplooidikten als maat voor de vetmassa, kan de gecorrigeerde arm- of beendiameter worden berekend. In dit onderzoek is gebruik gemaakt van de gecorrigeerde bovenarmdiameter berekend met behulp van de formule van Novak (Novak, 1963).

In dit onderzoek werden ter benadering van de spiermassa de volgende metingen verricht:

. omtrek van bovenarm, dijbeen en kuit (mm), met behulp van Martin's Circummètre;

. gecorrigeerde bovenarmdiameter (diameter van de bovenarm minus het onderhuidse vet ter plaatse).

Alle metingen die in ons onderzoek zijn toegepast, werden voor zover mogelijk aan de linkerzijde van het lichaam uitgevoerd, volgens de beschrijving van Tanner (Tanner, 1969).

De betrouwbaarheid van een aantal metingen is beschreven in hoofdstuk 6.1..

Verwachtingen

Op grond van de verzamelde gegevens spreken wij de verwachting uit dat onder invloed van twee extra lessen lichamelijke oefening per week:

. de skeletlengte zowel als de skeletbreedtematen zullen toenemen;

. de vetmassa zal afnemen;

. de spiermassa zal toenemen.

4.1.2. FYSIOLOGISCHE VARIABELEN

In het 'Voorstel Leerplan Rijksscholen Voor Lichamelijke Oefening voor V.W.O., H.A.V.O. en M.A.V.O. 1968' wordt onder meer de doelstelling van de lichamelijke oefening omschreven als: 'Het onderwijs in de

lichamelijke opvoeding heeft ten doel bij te dragen tot de vorming van de persoonlijkheid en is, met voortdurende inachtneming daarvan mede gericht op o.a. vergroting van de bereidheid en de geschiktheid tot het leveren van prestaties.'

Deze prestatieverhoging zou zowel op lichamelijk als op geestelijk gebied kunnen plaatsvinden (Rijsdorp, 1952).

Om veranderingen in het lichamelijk prestatievermogen te kunnen meten, is het allereerst noodzakelijk vast te stellen welke eigenschappen aan een lichamelijke prestatie ten grondslag liggen.

Theoretisch kunnen vanuit het gezichtspunt van de energievoorziening lichamelijke prestaties onderscheiden worden naar de duur van die prestaties (Shephard, 1968). In zijn algemeenheid kan men stellen dat bij alle prestaties die korter duren dan één tot twee minuten, de spieren voornamelijk van energie worden voorzien zonder dat daarbij zuurstof nodig is. Kort na het begin van de prestatie wordt alle energie uit energierijke fosfaten (ATP en CP), welke de spier in voorraad heeft, anaeroob vrijgemaakt. Al spoedig, na 5 - 10 seconden, wordt de energietoevoer tot stand gebracht door de splitsing van het in de spier aanwezige suiker dat onder vorming van melkzuur de energievoorziening steeds meer overneemt. Maar ook dit deel van de anaerobe energievoorziening is na enkele minuten uitgeput.

Het is iedereen evenwel bekend dat wij wel degelijk prestaties kunnen leveren die veel langer duren dan twee minuten. Enige tijd na het begin van de prestatie komt dan ook een andere energievoorziening langzamerhand op gang, zodat na vijf minuten bijna alle beschikbare energie van dat moment wordt vrijgemaakt door een oxydatie van suikers en vetten met behulp van zuurstof. Wanneer voldoende brandstoffen voorhanden zijn, is deze energievrijmaking in principe onuitputtelijk. De zuurstof kan namelijk bij een voldoende ventilatie en circulatie voortdurend naar de spieren getransporteerd worden. Aan brandstoffen is geen gebrek omdat van suikers en vetten een aanzienlijke energievoorraad in het lichaam zelf blijkt te zijn opgeslagen.

Vanuit dit energetische gezichtspunt zou men mogen veronderstellen dat er twee prestatiebepalende eigenschappen te onderscheiden zijn, te weten: 1) het anaerobe vermogen (anaerobic power) dat van belang is voor alle prestaties die korter duren dan 1 - 2 minuten, en 2) het aerobe vermogen (aerobic power) dat belangrijk is voor alle prestaties

4. AFHANKELIJKE VARIABELEN

4.1. ANTROPOMETRISCHE EN FYSIOLOGISCHE VARIABELEN

4.1.1. ANTROPOMETRISCHE VARIABELEN

De lichamelijke opvoeding stelt zich ten doel de lichamelijke ontwikkeling van de leerlingen gunstig te beïnvloeden. Om na te gaan in hoeverre die doelstellingen gerealiseerd worden wanneer zelfs twee lessen lichamelijke oefening aan het lesrooster worden toegevoegd, dient de lichaamsbouw meetbaar gemaakt te worden. Met behulp van antropometrische meetmethoden kan een indruk worden verkregen van de bijdrage van de diverse weefsels aan de lichaamsamenstelling.

Invloed van beweging op bouw en samenstelling van het lichaam

De bouw en samenstelling van het menselijk lichaam wordt voornamelijk bepaald door de hoeveelheid skelet-, spier- en vetweefsel. Alhoewel onder overigens 'normale' omstandigheden erfelijke factoren een overheersende rol spelen bij het bepalen van iemands algemene morfologische kenmerken, is bekend dat met name de ontwikkeling van spier- en vetweefsel wordt beïnvloed door uitwendige factoren zoals lichaamsbeweging en voeding. Aangezien wij in dit onderzoek te maken hebben met 12 en 13-jarigen, zal hun lichaamsbouw veranderingen ondergaan onder invloed van groeiprocessen. Het is dan ook niet uitgesloten dat de lichamelijke activiteiten zowel op de groei als op de ontwikkeling van de verschillende weefsels van invloed is.

Een minimum aan spierarbeid is een voorwaarde voor een normale groei van bot- en spierweefsel (Rarick, 1960; Malina, 1969). Hoe groot dat minimum moet zijn, is echter niet bekend. De bewegingsdrang van kinderen is over het algemeen echter dermate groot dat dit minimum ruimschoots zal worden overschreden.

Tevens is bekend dat vooral activiteiten met een hoge intensiteit een toename van de spierdoorsnede en de spierkracht veroorzaken (Rarick, 1960; Malina, 1969). Deze veranderingen zijn echter van tijdelijke aard en verdwijnen zodra men de training beëindigt. Gelijktijdig met deze toename in spierweefsel vindt een afname van de hoeveelheid vetweefsel plaats (Parizkova, 1963; Rarick, 1960; Malina, 1969). De oorzaak van deze verandering is gelegen in de verhouding energieopname en energieafgifte. Verhoogde lichamelijke activiteit veroorzaakt een

toename van de vetvrije massa ten koste van aanwezig vetweefsel. Het omgekeerde vindt plaats indien de mate van activiteit vermindert (Parizkova, 1963; Malina, 1969).

Ook bij onderzoek in Nederland naar de voedingstoestand van diverse bevolkingsgroepen constateerde men dat groepen met een grotere hoeveelheid dagelijkse lichamelijke activiteit minder lichaamsvet en meer spiermassa hebben dan de minder actieve groepen (Luyken, 1969 b en c). Bij een groep jongelui van 15 - 25 jaar werden deze veranderingen reeds waargenomen na afloop van een vierweekse cursus van zware lichamelijke activiteit (Luyken, 1969 a).

Op grond van bovenstaande gegevens kan in dit onderzoek verwacht worden dat bij meer lichamelijke oefening de vetmassa zal afnemen en de spiermassa zal toenemen.

Over de invloed van lichamelijke inspanning op het skelet is veel minder feitenmateriaal beschikbaar. Het ziet er echter naar uit dat lichamelijke inspanning tijdens de groeifase vooral de breedtegroei stimuleert (Rarick, 1960; Biener, 1961; Parizkova, 1963; Malina, 1969), terwijl in de lengtegroei weinig verandering optreedt (Rarick, 1960; Parizkova, 1963; Malina, 1969). Volgens Biener zal de lengtegroei door lichamelijke oefening alleen maar vertraagd worden, echter zonder dat de uiteindelijke lichaamslengte daardoor beïnvloed wordt. Dit effect treedt volgens Malina alleen op indien er grote druk op de botten wordt uitgeoefend. In de botstukken die niet betrokken zijn bij het dragen van het lichaam lijkt ook de lengtegroei door inspanning gestimuleerd te worden (Rarick, 1960; Malina, 1969).

Alhoewel de evidentie voor veranderingen in de groei van het skelet onder invloed van lichamelijke inspanning, met name de lengtegroei, niet groot is, verwachten wij toch een toename in de lengte zowel als in de breedtematen van het skelet, als gevolg van extra lessen lichamelijke oefening.

Antropometrische beoordeling van bouw en samenstelling

De basis voor een beoordeling van de lichaamsbouw wordt gevormd door het lichaamsgewicht en de lichaamslengte bij een gegeven leeftijd. Deze maten geven echter onvoldoende informatie.

Hoewel het lichaamsgewicht afhankelijk is van de verhouding waarin skelet-, spier- en vetweefsel voorkomen, geeft het geen informatie

over de afzonderlijke weefselmassa's. Vooral als men nagaat dat de variantie van het lichaamsgewicht van jonge volwassenen voor ongeveer 30% wordt verklaard uit de variabiliteit van spier- en vetmassa tezamen (Tanner, 1964), is duidelijk dat het lichaamsgewicht alléén een te grove maat vormt. Daar komt bij dat de vetmassa, maar vooral de spiermassa, op jongere leeftijd nog aan sterke veranderingen onderhevig is.

Bepaling skeletbouw

Als maat voor de bouw van het skelet geeft de lichaamslengte onvoldoende informatie. Factoranalytisch onderzoek heeft uitgewezen dat er behalve een factor skeletgrootte (lengte) een aantal factoren zijn die vormgeven aan de skeletbouw (Luyken, 1967). Terwijl de skeletgrootte wordt bepaald door lichaamslengte, romplengte of beenlengte, zijn het vooral de breedtematen zoals schouder-, borst- en bekkenbreedte, alsmede de botbreedte die de vorm aanbrengen. Vooral de breedte van de pijpbeenderen, gemeten op röntgenfoto's of als condylenwijdte, wordt beschouwd als de belangrijkste component van het skelet naast de skeletgrootte (Tanner, 1964). Deze breedtefactor verklaart ongeveer 10% van de variantie van alle lichaamsafmetingen, terwijl de skeletgrootte en botdikte tezamen 25-30% van de variantie voor hun rekening nemen (Tanner, 1964).

Ter benadering van de algehele groei en ontwikkeling werden in dit onderzoek zowel aan het begin als aan het einde van het schooljaar door getrainde personen de volgende metingen verricht:

- totaal lichaamsgewicht (kg) met behulp van een weegschaal van van Vucht's Concern, type Spido;
- lichaamslengte (mm), met behulp van een in het Laboratorium voor Psychofysiologie geconstrueerde digitaal afleesbare lengtemeter.

Ter benadering van de breedte van het skelet werden de volgende metingen verricht:

- schouder- en bekkenbreedte (mm), met behulp van een digitaal afleesbare meetlat, de Harpenden Antropometer;
- breedte van pols en knie (mm), met behulp van een Mauser schuifmaat, waarvan de bekranden waren verbreed.

Bepaling vetmassa

Voor het bepalen van de hoeveelheid vetmassa moet men zich het menselijk

lichaam voorstellen als de som van twee min of meer onafhankelijke compartimenten, namelijk de vetmassa en de vetvrije massa (v.v.m.). De vetmassa omvat al het lichaamsvet ongeacht de plaats en structuur. De vetvrije massa omvat de massa actief functionerend weefsel, te weten spier-, skeletmassa en de resterende weke delen, inclusief het bloed.

Om de hoeveelheid vetmassa of vetvrije massa te bepalen, kan gebruik gemaakt worden van een aantal methoden. De meest bekende methoden vormen de densitometrie en de antropometrie.

Bij de densitometrie berekent men de lichaamsdensiteit of -dichtheid door het lichaamsgewicht te delen door het volume. Het volume kan hierbij worden afgeleid door directe meting van de waterverplaatsing door het lichaam, of indirect via de bepaling van het lichaamsgewicht onder water. Aangezien de densiteit van spier-, skelet- en vetmassa verschilt, zal de totale lichaamsdichtheid uiteindelijk bepaald worden door de onderlinge verhouding van genoemde weefselsoorten. Via de meting van de densiteit kan men het aandeel vetmassa van het totale lichaamsgewicht berekenen (Brozek, 1963).

Aangezien de bepaling van de densiteit een gecompliceerde en kostbare methode vormt, heeft men getracht verbanden te vinden met een aantal antropometrische variabelen om zodoende de vetmassa of vetvrije massa met behulp van deze relatief eenvoudige metingen te kunnen schatten. Alhoewel men erin slaagde hoge correlaties te vinden tussen de densitometrie en de antropometrische voorspellingen van vetmassa en vetvrije massa, dient beseft te worden dat de hoge voorspellende waarde van de diverse formules alleen opgaat wanneer ze worden toegepast op eenzelfde populatie als waarop de formule oorspronkelijk was gebaseerd (Wilmore, 1969; Pieters, 1972). De onderwaterweegmethode of een vergelijkbare methode voor het meten van het lichaamsvolume met daarbij de directe berekening van het residuaire longvolume wordt algemeen als de meest nauwkeurige meting beschouwd.

In het vooronderzoek (Oud e.a., 1971) is bij 12 en 13-jarige jongens uit de brugklassen van het St. Ignatius College zowel de onderwaterweegmethode als de antropometrie toegepast. De regressievergelijking die uit deze gegevens werd opgesteld, maakte het mogelijk ditmaal te volstaan met antropometrische metingen. Uit twee huidplooien, triceps en subscapularis, kan het percentage vet worden geschat ($r = 0,87$) (Kemper, 1970, unpubl.).

In dit onderzoek werden ter benadering van de vetmassa de volgende metingen verricht:

. huidplooi van biceps, triceps, subscapularis en suprailiaca (mm), met behulp van een huidplooidiktemeter, de Harpenden Skinfold Caliper.

Bepaling spiermassa

De bijdrage van de spiermassa aan de lichaamssamenstelling kan eveneens langs indirecte weg met behulp van antropometrische metingen worden bepaald. Door gebruik te maken van omtreksmaten van armen of benen, als maat voor spier- en vetmassa tezamen, en huidplooidikten als maat voor de vetmassa, kan de gecorrigeerde arm- of beendiameter worden berekend. In dit onderzoek is gebruik gemaakt van de gecorrigeerde bovenarmdiameter berekend met behulp van de formule van Novak (Novak, 1963).

In dit onderzoek werden ter benadering van de spiermassa de volgende metingen verricht:

. omtrek van bovenarm, dijbeen en kuit (mm), met behulp van Martin's Circummètre;

. gecorrigeerde bovenarmdiameter (diameter van de bovenarm minus het onderhuidse vet ter plaatse).

Alle metingen die in ons onderzoek zijn toegepast, werden voor zover mogelijk aan de linkerzijde van het lichaam uitgevoerd, volgens de beschrijving van Tanner (Tanner, 1969).

De betrouwbaarheid van een aantal metingen is beschreven in hoofdstuk 6.1..

Verwachtingen

Op grond van de verzamelde gegevens spreken wij de verwachting uit dat onder invloed van twee extra lessen lichamelijke oefening per week:

. de skeletlengte zowel als de skeletbreedtematen zullen toenemen;

. de vetmassa zal afnemen;

. de spiermassa zal toenemen.

4.1.2. FYSIOLOGISCHE VARIABELEN

In het 'Voorstel Leerplan Rijksscholen Voor Lichamelijke Oefening voor V.W.O., H.A.V.O. en M.A.V.O. 1968' wordt onder meer de doelstelling van de lichamelijke oefening omschreven als: 'Het onderwijs in de

lichamelijke opvoeding heeft ten doel bij te dragen tot de vorming van de persoonlijkheid en is, met voortdurende inachtneming daarvan mede gericht op o.a. vergroting van de bereidheid en de geschiktheid tot het leveren van prestaties.'

Deze prestatieverhoging zou zowel op lichamelijk als op geestelijk gebied kunnen plaatsvinden (Rijsdorp, 1952).

Om veranderingen in het lichamelijk prestatievermogen te kunnen meten, is het allereerst noodzakelijk vast te stellen welke eigenschappen aan een lichamelijke prestatie ten grondslag liggen.

Theoretisch kunnen vanuit het gezichtspunt van de energievoorziening lichamelijke prestaties onderscheiden worden naar de duur van die prestaties (Shephard, 1968). In zijn algemeenheid kan men stellen dat bij alle prestaties die korter duren dan één tot twee minuten, de spieren voornamelijk van energie worden voorzien zonder dat daarbij zuurstof nodig is. Kort na het begin van de prestatie wordt alle energie uit energierijke fosfaten (ATP en CP), welke de spier in voorraad heeft, anaeroob vrijgemaakt. Al spoedig, na 5 - 10 seconden, wordt de energietoevoer tot stand gebracht door de splitsing van het in de spier aanwezige suiker dat onder vorming van melkzuur de energievoorziening steeds meer overneemt. Maar ook dit deel van de anaerobe energievoorziening is na enkele minuten uitgeput.

Het is iedereen evenwel bekend dat wij wel degelijk prestaties kunnen leveren die veel langer duren dan twee minuten. Enige tijd na het begin van de prestatie komt dan ook een andere energievoorziening langzamerhand op gang, zodat na vijf minuten bijna alle beschikbare energie van dat moment wordt vrijgemaakt door een oxydatie van suikers en vetten met behulp van zuurstof. Wanneer voldoende brandstoffen voorhanden zijn, is deze energievrijmaking in principe onuitputtelijk. De zuurstof kan namelijk bij een voldoende ventilatie en circulatie voortdurend naar de spieren getransporteerd worden. Aan brandstoffen is geen gebrek omdat van suikers en vetten een aanzienlijke energievoorraad in het lichaam zelf blijkt te zijn opgeslagen.

Vanuit dit energetische gezichtspunt zou men mogen veronderstellen dat er twee prestatiebepalende eigenschappen te onderscheiden zijn, te weten: 1) het anaerobe vermogen (anaerobic power) dat van belang is voor alle prestaties die korter duren dan 1 - 2 minuten, en 2) het aerobe vermogen (aerobic power) dat belangrijk is voor alle prestaties

die langer duren dan twee minuten (Asmussen, 1969).

Nu is de energievoorziening van de spieren lang niet altijd de belangrijkste grondeigenschap voor lichamelijke prestaties. Wanneer slechts een gering aantal spieren tegelijk maximaal, of wanneer een groot aantal spieren niet maximaal wordt belast, zullen de anaerobe en aerobe processen niet de beperkende factoren vormen. Vanuit fysiologisch standpunt spelen dan dikwijls geheel andere factoren een rol van betekenis. Deze factoren hebben alle te maken met het neuro-musculaire systeem (Astrand e.a., 1970). Onze skeletspieren kunnen slechts dan contraheren, wanneer ze geëxciteerd worden vanuit het centrale zenuwstelsel. En bij diverse sporten zal de prestatie niet zozeer beperkt worden door een tekort aan energie, maar door beperkingen die liggen in de eigenschappen van het neuro-musculaire systeem zelf, zoals de kracht of snelheid waarmee de spieren kunnen verkorten en de regulering via het centrale zenuwstelsel van ingewikkelde contractiepatronen, die afhangen van informatie via de zintuigen en ervaringsopslag.

Langs theoretische weg kan men zo het bestaan van een aantal andere prestatiebepalende eigenschappen aannemelijk maken. Deze eigenschappen krijgen in de praktijk een duidelijke betekenis wanneer men ze ook meetbaar kan maken. In dit verband zijn metingen of tests noodzakelijk. Dergelijke tests moeten dus in staat zijn om deze prestatiebepalende eigenschappen zo valide mogelijk te meten. Het resultaat van elke test moet dan ook zo weinig mogelijk door andere eigenschappen worden beïnvloed.

De meest bekende eigenschappen die worden onderscheiden, zijn: kracht, snelheid, lenigheid, coördinatie, behendigheid, evenwicht en uithoudingsvermogen (Fleishman, 1964). Deze opsomming is verre van uitputtend, maar ze geeft slechts aan welke veel genoemde eigenschappen men in de praktijk en theorie van de lichamelijke oefening en de sport tegenkomt. De meeste van deze eigenschappen worden ook nooit goed gedefinieerd. Men maakt duidelijk wat men met de eigenschap bedoelt door te vermelden bij welke tak van sport of bij welke prestatie die eigenschap van groot belang is.

Voor wat betreft de energetische aspecten van het lichamelijk prestatievermogen zal worden ingegaan op de meting van het aerobe vermogen met behulp van de fietsergometer. Tevens zullen eenvoudige prestatietests voor het meten van een aantal eigenschappen, waaronder aerobe

vermogen, kracht, snelheid en lenigheid, bekeken worden. Meting van het anaerobe vermogen zal achterwege gelaten worden in verband met het feit dat uit factoranalytische studies blijkt dat deze eigenschap niet is af te zonderen, aangezien de eigenschappen kracht en/of snelheid gelijktijdig een belangrijke rol spelen.

4.1.2.1. HET AEROBE VERMOGEN GEMETEN MET BEHULP VAN EEN FIETSERGOMETER-BELASTING

Onder het aerobe vermogen verstaat men de hoogste zuurstofopname uitgedrukt in liters per minuut, welke een individu kan bereiken bij lichamelijke arbeid tijdens het ademen van normale atmosferische lucht op zeeniveau (Astrand e.a., 1970).

Omdat de belasting tijdens de dagelijkse lichamelijke activiteit meestal afhankelijk is van het lichaamsgewicht, is het wenselijk de zuurstofopname ook per kilogram lichaamsgewicht uit te drukken. Over het algemeen wordt de maximale zuurstofopname als de beste maat beschouwd voor het aerobe vermogen (Asmussen, 1969; Shephard, 1968). De maximale zuurstofopname ($\dot{V}O_2max$) is echter alleen in speciaal ingerichte laboratoria tijdens gestandariseerde belastingsproeven vast te stellen.

In het laboratorium kan de $\dot{V}O_2max$ direct gemeten worden met behulp van een maximaaltest. Daarbij wordt van de proefpersoon een 'maximale' belasting gevraagd. Omdat dit bij jeugdige personen lang niet altijd mogelijk is, worden ook wel submaximaaltests toegepast. Hierbij wordt de $\dot{V}O_2max$ indirect gemeten door deze te voorspellen uit een of meerdere metingen van de zuurstofopname en de hartfrequentie bij submaximale belastingen.

Keuze van het belastingsprogramma

Het International Biological Programma (I.B.P.) heeft voorgesteld om de submaximaaltests uit te voeren op een fietsergometer. De voorgestelde test heeft een totale duur van twaalf minuten en kan verdeeld worden over vier periodes van elk drie minuten continue belasting. In plaats van een '4 x 3 minuten'-schema kan men ook kiezen voor een '3 x 4', of een '2 x 6'-schema (Weiner e.a., 1969).

Bij kinderen is het gewenst om de belasting op de fietsergometer in verband te brengen met het lichaamsgewicht; de verschillende belastingen kunnen dan worden uitgedrukt in kgm/kg.min. of in watt/kg..

Aangetoond kon worden dat bij 12 en 13-jarige jongens tijdens belastingen per kilogram lichaamsgewicht geen significante verschillen bestaan in gemiddelde hartfrequentie tussen zes verschillende gewichtsgroepen (Kemper, 1972).

Welke belasting men moet kiezen kan mede gebaseerd worden op de hartfrequentie die bereikt wordt in de voorafgaande belasting. Wanneer voor de eerste belastingsperiode een belasting van 1 watt/kg wordt gekozen, kan men ernaar streven dat alle volgende belastingsperiodes hartfrequenties opleveren tussen 110 en 180 slagen per minuut. Immers binnen dit traject is de relatie tussen hartfrequentie en zuurstofopname nagenoeg lineair (Mocellin e.a., 1970).

In schema 4.1. is de bij dit onderzoek gehanteerde belastingsprocedure uitgewerkt. Wij hebben daarbij gekozen voor een '4 x 3 minuten'-schema, omdat in voorgaand onderzoek bij eenzelfde populatie bleek dat bij een '3 x 6 minuten'-schema de grootste stijging van de hartfrequentie gedurende de eerste twee minuten optrad; gedurende de volgende vier minuten bleef de hartfrequentie gelijkmatig oplopen zonder dat een duidelijke steady-state aantoonbaar was (Kemper, 1972).

period	time	intensity			
1	3 min	1.0 Watt/kg choice of following load depends on mean heart rate (f_h) in the last minute of the preceding load			
		$f_{h=} \geq 120$		$f_h < 120$	
2	3 min	1.5 Watt/kg		2.0 Watt/kg	
		$f_{h=} \geq 140$	$f_h < 140$	$f_{h=} \geq 140$	$f_h < 140$
3	3 min	2.0 Watt/kg	2.5 Watt/kg	2.5 Watt/kg	3.0 Watt/kg
		$f_{h=} \geq 160$ — $f_h < 160$	$f_{h=} \geq 160$ — $f_h < 160$	$f_{h=} \geq 160$ — $f_h < 160$	$f_{h=} \geq 160$ — $f_h < 160$
4	3 min	2.5 Watt/kg — 3.0 Watt/kg	3.0 Watt/kg — 3.5 Watt/kg	3.0 Watt/kg — 3.5 Watt/kg	3.5 Watt/kg — 4.0 Watt/kg

Schema 4.1.: Procedure of the loading pattern on the bicycle ergometer.

'Physical Working Capacity' (W_{170})

Omdat het in ons onderzoek om technische en organisatorische redenen niet mogelijk bleek om betrouwbare zuurstofopnamemetingen te verrichten, hebben wij er de voorkeur aan gegeven om in plaats van de zuurstofopname ter bepaling van het aerobe vermogen, de 'Physical Working Capacity' (W_{170}) te meten.

De W_{170} is de belasting welke gepaard gaat met een hartfrequentie van 170 slagen per minuut (Wahlund, 1948) en deze is een betrouwbare maat voor het bepalen van het aerobe vermogen (Edholm, 1969).

Mellerowicz (1965) stelde voor om de hartfrequentie te vergelijken bij belastingen per kilogram lichaamsgewicht. De gemeten W_{170} wordt dan uitgedrukt in een vermogen per kilogram lichaamsgewicht (watt/kg).

Theoretisch gesproken is elk instrument dat ons mogelijk maakt om goed gedefinieerde belastingen aan te bieden geschikt, mits een groot percentage van de totale spiermassa wordt aangesproken.

Door de W.H.O. is in 1968 een overzicht gegeven van de voor- en nadelen die verbonden zijn aan de drie meest gebruikte instrumenten: de tredmolen, het opstapbankje en de fietsergometer. Bij de door ons gebruikte submaximale test ter meting van de W_{170} gaat de voorkeur uit naar de fietsergometer. Dit hangt samen met de eenvoud van de te verrichten arbeid: hoe gewoner de omstandigheden zijn waaronder men de test kan afnemen des te minder wordt de hartfrequentie beïnvloed. Dit zogenaamde habituatie effect levert bij twee opeenvolgende dagen van testen bij de proefpersoon een gemiddelde daling op in hartfrequentie; bij de tredmolen van 6 slagen/min, bij het opstapbankje van 4 slagen/min en bij de fietsergometer geen enkele daling (Shephard, 1969).

Omdat het hier gaat om een longitudinaal onderzoek waarbij tussen voor- en natest een trainingseffect moet worden aangetoond in de W_{170}, is het uiteraard van belang dat het habituatie effect zo gering mogelijk is.

De in ons onderzoek gemeten W_{170}-waarden blijken dus sterk afhankelijk van zowel de soort van belasting - fietsen, lopen of traplopen - als van verschillen in het belastingsschema. Bij vergelijking van de resultaten met ander onderzoek dient daarmee ernstig rekening gehouden te worden (Mocellin e.a., 1971; Placheta, 1971).

Fietsergometer

De fietsergometer belastingsproef werd uitgevoerd op het St. Ignatius College gedurende de normale schooluren in een aparte ruimte naast de gymnastiekzalen. De fietsergometer (fabrikaat Lode) werd elektrisch geremd. Op regelmatige tijden werd de ergometer gecontroleerd en afgeregeld (Bleeker e.a.,1960).

Door wijziging in stuurafstand en zadelhoogte werd deze fietsergometer geschikt gemaakt voor het testen van proefpersonen met grote verschillen in lengte. De meter, welke de te trappen belasting aangeeft, was voor hen niet zichtbaar. Om gedurende de fietsproef een constante trapfrequentie van ongeveer 60 omw/min te kunnen aanhouden, was een voor de proefpersoon duidelijk zichtbare toerenteller aangebracht.

Proefpersonen, registratie en belastingsprocedure

Vooraf werden de proefpersonen gewogen in sportbroekje en gymnastiekschoenen. Dit gewicht, afgerond op 5 kg, werd gebruikt voor de zwaarte van de in te stellen belasting.

Terwijl de proefpersoon op de fiets zat, werden op de borstwand de E.C.G.-electrodes aangebracht, daarna een bloeddrukmanchet om de rechter bovenarm, en ten slotte een ademmasker over mond en neus.

De belastingsproef werd voorafgegaan door registraties in rust tijdens drie minuten zitten op de fiets. Daarna begon de proefpersoon te trappen tegen een belasting van 1 watt per kilogram lichaamsgewicht. Gedurende de drie volgende belastingsperiodes van elk drie minuten werd de belasting verhoogd, afhankelijk van de gemiddelde hartfrequentie, gemeten in de laatste minuut van de voorafgaande belastingsperiode.

De belastingsproef werd beëindigd met een herstelperiode van vijf minuten. Gedurende elke minuut werd continu op een oscilloscoop het precordiale E.C.G. weergegeven. Door een cardiotachometer (fabrikaat Rood) werd het R-R.-interval omgezet in een hartfrequentie en vervolgens geschreven op een schrijver (fabrikaat Servogor).

In de laatste minuten van elke periode werden langs indirecte weg de systolische en diastolische bloeddruk bepaald (Autotest).

De uitgeademde lucht van de proefpersoon werd continu geanalyseerd

op het percentage zuurstof; met behulp van net gemeten ademminutenvolume werd de zuurstofopname per minuut berekend (app. type Oxycon) in STPD-waarden, d.w.z. gecorrigeerd naar een standaardtemperatuur ($0^{o}C$), druk (760 mmHg) en droogte.

De W_{170} van elke proefpersoon werd vervolgens bepaald uit een regressielijn die berekend wordt met de gemeten hartfrequenties in de laatste minuut van elke belastingsperiode. Waarden beneden de 120 en boven de 180 sl/min werden om lineariteitsredenen buiten beschouwing gelaten. Omdat de W_{170} is uitgedrukt in een wattage per kilogram lichaamsgewicht kan een verandering van de W_{170} van een proefpersoon op twee manieren tot stand komen: door een vergroting van het aerobe vermogen van de proefpersoon zelf, maar ook door een verandering van het lichaamsgewicht.

Met andere woorden, wanneer als gevolg van de onafhankelijke variabele de jongens uit de experimentele groep een groter vermogen op de fiets blijken te trappen, neemt de W_{170} toe; anderzijds kan bij een gelijkblijvend vermogen de W_{170} eveneens toenemen als tegelijkertijd het lichaamsgewicht is gedaald tengevolge van afname van de vetmassa.

Verwachting

Wij spreken de verwachting uit dat als gevolg van de extra lessen de W_{170}(watt/kg)bij de experimentele groep zal toenemen ten opzichte van de controlegroep.

4.1.2.2. HET AEROBE VERMOGEN GEMETEN MET BEHULP VAN EENVOUDIGE PRESTATIETESTS

Aangezien de meting van het aerobe vermogen op grote schaal met behulp van de fietsergometer moeilijk te realiseren is, heeft men getracht veldtests te ontwikkelen met de bedoeling toch een betrouwbare maat voor het aerobe vermogen te verkrijgen (Brouha, 1943; Balke, 1963).

Vast staat dat de meting van het maximale zuurstofopnamevermogen een betrouwbare maat vormt voor het aerobe vermogen. Men heeft daarom de samenhang bestudeerd tussen de resultaten van de laboratoriumtests en de diverse veldtests, waarvan de step-tests (Brouha, 1943; Master, 1968) en de looptests (Balke, 1963; AAHPER, 1967; CAHPER, 1967; Cooper, 1968) de meest bekende zijn.

Samenhang tussen veld- en laboratoriumtests

Hollmann (1966) onderzocht de mate van samenhang van een aantal steptests met de $\dot{V}O_2$-max., gemeten tijdens een maximaaltest. Bij geen enkele van deze tests bleek een significante correlatie met de $\dot{V}O_2$-max. aangetoond te kunnen worden; de 'beste' samenhang werd over het algemeen gevonden met die tests waarvan de belastingsduur en de belastingsintensiteit het grootst was.

Rasch e.a. (1964) vergeleken het uithoudingsvermogen van militairen - gemeten als loopprestatie op de drie mijl - met de resultaten van een aantal physical fitness tests. Alleen Balke's tredmolentest kon een goede voorspelling geven van de loopprestatie.

Alhoewel Simons e.a. (1970) looptests in hun batterij hadden opgenomen die verondersteld werden het aerobe vermogen te meten, slaagden zij er niet in deze eigenschap als een aparte factor in de factoranalyse af te zonderen. Opvallend daarbij was dat deze tests ladingen vertoonden op de factor loopsnelheid. Aangezien het afleggen van deze tests, de 160 en 240-meter shuttle run, niet veel meer dan één á anderhalve minuut in beslag zal nemen, is het onmogelijk om vanuit energetisch gezichtspunt iets te kunnen meten van het aerobe vermogen. Immers voor het meten van het aerobe vermogen dient de arbeid tenminste twee á drie minuten te duren (Shephard, 1968).

Twee zeer veel toegepaste testbatterijen, de A.A.H.P.E.R. Youth Fitness Test (A.A.H.P.E.R., 1967), en de C.A.H.P.E.R. Physical Performance Test (C.A.H.P.E.R., 1967) gebruiken als maat voor het aerobe vermogen respectievelijk een loop over een afstand van zeshonderd en driehonderd yards. Ook hiervoor geldt dat men bij het afleggen van deze tests nauwelijks een belangrijk aandeel in de energievoorziening door aerobe oxydatie mag verwachten (Cumming, 1971). Bij kinderen en volwassenen zal de inspanning pas bij een arbeidsduur van meer dan tien minuten voor meer dan tachtig procent aeroob verlopen (Hansen, 1967).

Correlation between running performance and $\dot{V}O_2$-max.			
distance	age	corr.coeff.	author(s)
300 yards	6 - 17	-.65	Cumming e.a. (1967)
440 "	18 - 25	-.22	Wiley e.a. (1972)
600 "	12 - 13	-.66	Metz e.a. (1970)
600 "	14 - 15	-.27	Metz e.a. (1970)
600 "	20	-.64	Falls e.a. (1966)
600 "	14 - 15	-.62	Doolittle e.a. (1968)
600 "	16 - 17	-.53	Olree e.a. (1965)
600 "	16 - 59	-.27	Drake e.a. (1968)
1 mile	18 - 25	-.29	Wiley e.a. (1972)
2 miles	18 - 25	-.47	Wiley e.a. (1972)
12 minutes	14 - 15	.90	Doolittle e.a. (1968)
12 "	11 - 14	.65	Maksud e.a. (1971)
12 "	adults	.90	Cooper (1968)
12 "	8 - 17	.53	Cumming (1973)

Tabel 4.2. : Correlation coefficients between $\dot{V}O_2$-max and running performances at different ages; data given by different authors.

In table 4.2. zijn door diverse onderzoekers correlaties vermeld tussen looprestaties en het maximale zuurstofopnamevermogen. Met het toenemen van de loopafstand blijkt nauwelijks een tendens aanwezig tot een betere samenhang met de $\dot{V}O_2$-max..

Uit deze correlatiecoëfficienten blijkt echter wel dat van de bestaande looptests de prestaties op de 12-minutenloop over het algemeen de beste relatie vertonen met de meting van de $\dot{V}O_2$-max.. Ook de test-hertestbetrouwbaarheid steunt de keuze van de 12-minutenloop als eenvoudige test voor het meten van het aerobe vermogen van 12 en 13-jarige jongens (tabel 4.3.).

reliability 12-min. run walk		
age (years)	corr.coeff.	author(s)
11 - 14	.92	Maksud e.a. (1971)
12	.66	Bovend'eerdt e.a.(1973)
13	. 77	Bovend'eerdt e.a.(1973)
14	.73	Bovend'eerdt e.a.(1973)
14 - 15	.94	Doolittle e.a. (1968)
15	.81	Bovend'eerdt e.a.(1973)
16	.70	Bovend'eerdt e.a.(1973
17	.88	Bovend'eerdt e.a.(1973)
adults	.94	Cooper (1968)

Tabel 4.3. : Test-retest reliability of 12-min. run walk, data given by different authors.

Ondanks voornoemde correlatie tussen 12-minutenloop en $\dot{V}O_2$-max., en de betrouwbaarheid van deze looptest, ziet het ernaar uit dat deze test de laboratoriummetingen niet kan vervangen. De reden is dat, hoewel de prestatie op de 12-minutenloop soms een vrij hoge correlatie met de maximale zuurstofopname van 0,90 vertoont (Cooper, 1968; Doolittle, 1968), deze toch een matige voorspeller blijkt te zijn van de $\dot{V}O_2$-max.

Bij Cooper (1968) blijken personen met eenzelfde prestatie op de 12-minutenloop (bijvoorbeeld 1,6 mijl) zowel een zeer hoge $\dot{V}O_2$-max. (52 ml/kg/min) als een betrekkelijke lage $\dot{V}O_2$-max. (40 ml/kg/min) te kunnen hebben. Het verschil tussen de werkelijke $\dot{V}O_2$-max. en de voorspelde $\dot{V}O_2$-max. uit de prestatie op de 12-minutenloop blijkt 10 - 15% te bedragen.

Tevens valt op dat bij sommige onderzoekingen de test voor het meten van het aerobe vermogen even hoog of zelfs lager correleert met de $\dot{V}O_2$-max. dan andere physical fitness test items (Cumming e.a., 1967; Metz e.a., 1970).

De verklaring voor een lage correlatie wordt dan gezocht in het

feit dat men met proefpersonen te maken heeft die wat betreft hun groei en ontwikkeling sterk uiteenlopen (Maksud, 1971; Metz, 1970). Ook de motivatie van de proefpersonen zou een verklaring kunnen betekenen (Metz, 1970). Cumming (1973) tenslotte, geeft aan dat bij 12 tot 17 jaar oude jongens en meisjes het resultaat op de 12-minutenloop een goede voorspeller kan zijn van de $\dot{V}O_2$-max., zelfs even goed - zo niet beter - dan de voorspelling met behulp van de fietsergometer. Hij voegt daar echter de waarschuwing aan toe dat dit alleen het geval is wanneer we te maken hebben met goed gemotiveerde jongeren.

Testkeuze en meetprocedure

De 12-minutes run walk test (Cooper, 1968) is in dit onderzoek opgenomen als eenvoudige test voor het meten van het aerobe prestatievermogen. Aangezien dit aspect tevens is gemeten met behulp van de fietsergometerproef, zullen wij de resultaten van deze metingen vergelijken teneinde vast te stellen welke maat gebruikt zal worden voor de verdere verwerking van de testgegevens (zie factoranalyse 6.2.).

De 12-minutenloop werd per klas afgenomen door de betreffende leraar lichamelijke oefening. Daartoe werd op een tegelplateau een traject van 150 meter uitgezet, dat met behulp van genummerde verkeerskegels was onderverdeeld in stukken van elk 25 meter.

Voor het afnemen van de test werd de klas opgesplitst in twee groepen. Terwijl de ene helft van de klas de test aflegde, hielden de leerlingen in de andere groep het aantal afgelegde ronden bij van een bepaalde klasgenoot. Tevens werd na twaalf minuten vastgesteld welk deel van een ronde, aangegeven door de kegels, nog extra was afgelegd. Uit deze score kon de totaal afgelegde afstand worden berekend.

Verwachting

Verwacht wordt dat de prestaties op de 12-minutenloop, uitgedrukt in de afgelegde afstand, onder invloed van extra lessen lichamelijke oefening zullen toenemen.

4.1.2.3. HET METEN VAN DE EIGENSCHAPPEN VAN HET NEURO-MUSCULAIRE SYSTEEM MET BEHULP VAN EENVOUDIGE PRESTATIETESTS

Door het ontbreken van een betrouwbare internationale maat voor het meten van de overige veronderstelde lichamelijke eigenschappen, is

het aantal uit de praktijk voortkomende tests zeer groot.

Diverse onderzoekers hebben getracht met behulp van factoranalytische studies inzicht te verkrijgen in het overvloedige testmateriaal.

De meest uitgebreide onderzoekingen op dit gebied zijn uitgevoerd door Fleishman (1964). Hij analyseerde 30 tests voor het meten van de eigenschap kracht (strength), en 30 tests voor het meten van de eigenschappen snelheid(speed), lenigheid (flexibility), evenwicht (balance) en coördinatie. Tevens werden test-hertest betrouwbaarheidsmetingen uitgevoerd voor de strenth tests.

Doel van de onderzoeken van Fleishman was na te gaan in hoeverre de eigenschappen van het lichamelijk prestatievermogen die in de literatuur worden onderscheiden, aangetoond kunnen worden door de tests die experimenteel ontwikkeld zijn om die eigenschappen te meten (validiteit).

Uit de analyse van de strength tests kwamen de factoren 'dynamic', 'static', 'explosive' en 'trunk' strength naar voren. Analyse van de tests voor speed, flexibility, balance en coordination leverde de factoren 'dynamic' en 'extent flexibility', 'speed of limb movement', 'gross body equilibrium' en 'balance with visual cues' op.

De factorlading die elk van de bij het onderzoek betrokken testitems krijgt, maakt tevens mogelijk om vast te stellen met welke test een bepaalde factor het best kan worden gemeten. Voor het maken van deze testkeuze heeft men echter nog informatie nodig over de testhertest betrouwbaarheid, die bij Fleishman's onderzoek alleen is uitgevoerd over de strength tests.

In navolging van het werk van Fleishman hebben Simons e.a. (1970) onderzocht welke motorische basiseigenschappen bij jongens in de leeftijd van 12 tot 19 jaar te onderscheiden zijn. Met behulp van factoranalyse werd nagegaan in hoeverre de door Fleishman (1964) gevonden factoren ook bij deze jongere populatie konden worden aangetoond.

Hiertoe werden naast een groot aantal reeds door Fleishman (1964) opgenomen tests, tevens tests voor het meten van het cardio-respiratoire uithoudingsvermogen in het onderzoek opgenomen. Ten slotte werd van alle tests de test-hertestbetrouwbaarheid bepaald.

De resultaten van dit onderzoek geven aan dat bij verschillende leeftijdscategorieën andere basisvaardigheden naar voren kunnen komen. Simons e.a. (1970) onderscheiden op basis van zes factoranalyses, zeven factoren op het gebied van de motorische vaardigheden die zij als volgt interpreteren: 1) functionele-, 2) statische-, 3) explosieve- en 4) rompkracht, 5) lenigheid, 6) loopsnelheid en 7) snelheid van de ledematen.

Behalve dat zij het bestaan van deze zeven factoren hebben aangetoond, wordt het bestaan van twee andere eigenschappen vermoed, te weten: 8) lichaamsevenwicht en 9) oog-ledematen-coördinatie.

Opvallend in het onderzoek van Simons e.a. (1970) is dat bij de 12-jarige jongens de factor 'explosieve kracht' niet kon worden afgezonderd, terwijl de tests die deze eigenschap meten, duidelijke ladingen vertoonden op de factoren functionele en statische kracht, alsmede loopsnelheid. Tevens kon, bij zowel de 12 als de 13-jarigen, de factor 'rompkracht' niet worden onderscheiden.

'Twee factoren scheiden zich duidelijk af,de functionele en de statische kracht. Met het ouder worden schijnen deze factoren zich echter nog verder te differentiëren en dient er nog een onderscheid gemaakt te worden tussen explosieve kracht en een vorm van rompkracht' (Willems e.a., 1971).

Testkeuze

De keuze van de tests die in onderzoeken zijn aangewend voor het meten van veranderingen in het lichamelijk prestatievermogen, is voornamelijk gebaseerd op de resultaten van het onderzoek van Simons, aangezien dit het enige onderzoek is dat op grote schaal factoranalyse heeft toegepast op tests afgenomen bij een overeenkomstige proefgroep.

Voor het meten van de eigenschappen die bij 12 en 13-jarige jongens werden aangetoond, viel onze keuze op de test die op genoemde factor de hoogste lading vertoonde en bovendien een grote test-hertestbetrouwbaarheid had.

Voor een uitvoerige beschrijving van de in tabel 4.4. opgesomde tests verwijzen wij naar bijlage II.

Battery of simple performance tests					
factor	test	factor loading		test-retest r.	
		12 y.	13 y.	12 y.	13 y.
functional strength	BENT ARM HANG	.70	.70	.80	.80
static strength	HANDGRIP	.76	.80	.89	.87
explosive strength	VERTICAL JUMP except for 12 y.	--	.32	--	.83
trunk strength	except for 12 and 13 year	--	--	--	--
flexibility	SIT AND REACH	.55	.65	.92	.92
running speed	50 m SHUTTLE RUN	.78	.85	.68	.66
speed of limb movement	PLATE TAPPING	.56	.66	--	.60

Tabel 4.4. : Battery of simple performance tests. The choice of the specific tests have been made on the base of the results of Simons e.a. (1970); factor loadings as well as test-retest correlations are also from Simons e.a. (1970).

- De 'bent arm hang' meet hoelang iemand met gebogen armen aan een rekstok kan blijven hangen, terwijl zijn ogen zich ter hoogte van de stok moeten bevinden.
- De 'handgrip', gemeten met de Bettendorff dynamometer, bepaalt de handknijpkracht.
- Bij de 'vertical jump' wordt de relatieve spronghoogte bepaald door de absolute spronghoogte te verminderen met de reikhoogte. De meetlat werd ontwikkeld aan het Laboratorium voor Psychofysiologie.
- 'Sit and reach' meet de reikafstand in langzit. De gebruikte voetensteun annex meetlat is ontwikkeld door het Laboratorium voor Psychofysiologie.
- Bij de '50 meter shuttle run' wordt de score bepaald door de tijd nodig om een afstand tussen twee evenwijdige lijnen van tien meter vijf maal af te leggen.
- Bij 'plate tapping' ten slotte, wordt bepaald hoeveel tijd de proefpersoon nodig heeft om, met zijn voorkeurhand, afwisselend twee

schijven elk veertig maal aan te raken. De schijven, tachtig cm uit elkaar liggend, worden via een verstelbaar plateau op heuphoogte gebracht (Laboratorium voor Psychofysiologie).

Verwachtingen

De prestaties op de eenvoudige prestatietests zullen onder invloed van de extra lessen lichamelijke oefening meer toenemen bij de experimentele groep dan bij de controlegroep.

Dit betekent dat wij verwachten dat de scores op de 'bent arm hang', 'handgrip', 'vertical jump' en 'sit and reach' meer zullen toenemen bij de experimentele, dan bij de controlegroep, terwijl de scores op de 'shuttle run' en 'plate tapping' juist meer zullen afnemen.

4.1.2.4. HET METEN VAN DE VENTILATIEPRESTATIE

Het zuurstoftransport naar de spieren speelt bij vele lichamelijke prestaties een rol van betekenis (zie: fysiologische variabelen 4.1.2.). Dit transport vindt plaats door middel van ventilatie en circulatie, en wordt in ons onderzoek getest met behulp van de fietsergometer.

Tevens is gezocht naar een eenvoudige maat voor het functioneren van de ventilatie. Alhoewel de Vitale Capaciteit, het volume lucht dat na een maximale inademing maximaal kan worden uitgeademd, veelvuldig wordt gehanteerd als maat voor de ventilatieprestatie, kan door het ontbreken van het tijdselement niet veel waarde aan deze maat worden toegekend (Bass, 1962). Een betere indruk wordt verkregen door de éénsecondewaarde (FEV1), de vijfsecondewaarde (FEV5) en het ventilatiequotient (FEV %), waarbij de tijd wel een rol speelt (FEV = Forced Expired Volume). De invloed van de tijd is bij de FEV5 echter zo klein dat deze maat sterk overeenkomt met de Vitale Capaciteit (Bonjer, 1965). Dit geldt in mindere mate voor de FEV1, alhoewel het ontbreken van informatie over het longvolume vergelijking van éénsecondewaarden twijfelachtig maakt. De beste maat voor het vergelijken van de ventilatieprestatie van verschillende personen wordt dan ook gevormd door de FEV %, berekend uit het quotient van FEV en FEV5, vermenigvuldigd met 100.

In een onderzoek naar het effect van zes weken circuittraining tijdens de lessen lichamelijke oefening van 12 en 13-jarige jongens

van het St. Ignatius College werd een verbetering gevonden van de FEV1 en de FEV % (Kemper, 1970).

Ter benadering van de prestatie van de ademhaling werd gebruik gemaakt van: F.E.V.% (FEV1/FEV5 x 100), gemeten met een Vitalograph.

Verwachting

Onze verwachting is dat onder invloed van extra lessen lichamelijke oefening de ventilatieprestatie, gemeten als FEV %, bij de experimentele groep zal toenemen ten opzichte van de controlegroep.

4.1.3. GYMNOLOGISCHE VARIABELEN

Vooronderzoek

In het vooronderzoek dat in het schooljaar 1969/1970 op dezelfde school en eveneens met leerlingen uit de brugklassen werd uitgevoerd (Oud e.a., 1971), kon worden aangetoond dat als gevolg van de twee extra lessen lichamelijke oefening per week aan het einde van het schooljaar het cijfer voor lichamelijke oefening bij de experimentele groep hoger was geworden ten opzichte van de controlegroep. Blijkens een kort bericht in Folia Civitatis (1971), het blad van de Universiteit van Amsterdam, werden de resultaten van dit vooronderzoek als volgt samengevat: 'Wie op school meer lessen gymnastiek krijgt dan normaal, wordt beter in gymnastiek'.

In een eigen discussie naar aanleiding van deze resultaten (Kemper e.a., 1972) werd verondersteld dat deze verbeteringen voornamelijk tot stand zijn gekomen door een verbetering in neuro-musculaire coördinatie, zonder dat er sprake was van een duidelijke verandering in morfologie. De tests, welke gebruikt werden om de vorderingen in de lichamelijke oefening te meten, hadden dan ook sterk het karakter van motorische vaardigheden in aansluiting aan de praktijk van de lichamelijke oefening, waarin de motorische vaardigheden op de voorgrond staan. Dat het realiseren van dit 'nabije' doel wél blijkt te lukken, heeft waarschijnlijk als belangrijkste oorzaak dat juist dit doel de leraar zeer helder voor ogen staat. Bovendien kan de leraar ervoor zorgen dat bij zijn leerlingen dit doel eveneens concreet wordt.

Concretisering van de andere doelen in de lichamelijke oefening, zoals onder meer geformuleerd in het Voorstel leerplan Rijksscholen

(1968) heeft nauwelijks plaatsgevonden. Voor de leraar zijn doelstellingen, zoals vergroting van de bereidheid en geschiktheid tot het leveren van prestaties, de bevordering van de zin voor samenwerking en dergelijke, niet eenvoudig te concretiseren. Wanneer dit soort 'verre' doelen door de leraar nauwelijks geconcretiseerd kan worden, zal men niet mogen verwachten dat deze dan ook snel zullen worden bereikt (Westerhof, 1972). Dit is wellicht de reden dat de leraar lichamelijke oefening het grootste deel van de les primair gericht is op het aanleren of verbeteren van zuiver motorische vaardigheden of 'skills'. Zijn opleiding is er voor een groot deel op gericht om methodieken te leren waarmee bepaalde vaardigheden zo snel mogelijk kunnen worden bereikt.

Op grond van het bovenstaande kan men begrijpen waarom de gymnastiekprestaties wel verbeterd werden en andere lichamelijke en geestelijke kenmerken niet.

Ook uit het overzicht van onderzoekingen naar het effect van extra lessen lichamelijke oefening blijkt over het algemeen dat, wanneer er een effect kan worden aangetoond, het de gymnastiekprestaties zijn die verbeterd zijn (zie schema 1.1.).

Voor het onderzoek van Paschen (1971) naar het effect van een dagelijkse les lichamelijke oefening bij jongens en meisjes van de eerste en tweede klas van de basisschool is voor het eerst een testbatterij ontworpen waarmee specifieke gymnastiekprestaties konden worden onderzocht. Deze 'Bewegungsleistungtest' werd driemaal afgenomen en bestond uit de volgende onderdelen: balanceren, touwklimmen, boksspringen, springen over hindernissen, optrekken aan de rekstok, 40-meterloop, afstandsworp met een slagbal, doelen op een basket en springen over matten. Helaas bleek deze testbatterij aan het einde van de periode voor verschillende leerlingen te eenvoudig, zodat van hen geen vorderingen gemeten konden worden.

Gymnologische tests

Evenals in het vooronderzoek werden de vorderingen in de lichamelijke oefening gemeten met objectieve prestatietests, hier kortweg aangeduid met gymnologische tests, om deze te onderscheiden van de prestatietests welke beschreven zijn in 4.1.3..

Uitgaande van het leerplan werd voor de brugklassen een reeks

tests gebruikt die min of meer aansluit bij de geboden oefenstof in de voorafgaande lessen (zie 3. en bijlage II). Om het eerder genoemde bezwaar van de tests van Paschen te vermijden, maar bovendien om telkens met specifieke tests te kunnen aansluiten bij datgene wat op dat moment geoefend wordt, werden voor elke periode andere tests gebruikt. Voor de zwemlessen kon volstaan worden met dezelfde tests, voor elke periode twee.

De door deze gymnologische tests gemeten prestatie kon bestaan uit de volgende objectieve gegevens:

- de tijd die men nodig heeft om een gegeven beoefening uit te voeren;
- het aantal herhalingen dat men kan uitvoeren, al dan niet binnen een voorgeschreven tijdsduur;
- de afstand die men met een bepaalde oefening kan afleggen;
- een leergang van een bepaalde eindvorm, waarbij elke score een stap vormt in deze leergang.

Deze tests waren door de leraren lichamelijke oefening zelf geconstrueerd gedurende de jaren 1968 en 1969. In de twee daarop volgende jaren, 1970 en 1971, werden de criteria van elke test bij groepen van vergelijkbare brugklasleerlingen beproefd.

Elke test werd gescoord op een 5-puntsschaal, waarbij een nul een 'slechte' en een vier een 'uitstekende' prestatie op dat onderdeel betekende.

Alle tests werden zodanig aangepast dat de gemiddelde testscore, behaald door alle brugklasleerlingen, tussen de twee en drie lag. Tot dusver heeft geen validering van deze gymnologische tests plaatsgevonden. Het afnemen van de tests gebeurde door de leraar lichamelijke oefening zelf in zijn eigen klas, en in dezelfde lessen op een zodanig tijdstip, dat de betreffende eindvormen van het leerplan reeds één of meerdere malen in voorafgaande lessen aan de orde waren geweest.

In totaal werden gedurende het schooljaar 23 gymnastiektests en 8 zwemtests afgenomen.

Om een indruk te kunnen krijgen van het verloop van de vorderingen van de experimentele en de controlegroep werd het schooljaar in vier perioden verdeeld: 1. tot de herfstvakantie; 2. tot de kerstvakantie; 3. tot de paasvakantie; 4. tot de zomervakantie.

Tabel 4.5. geeft de verdeling weer van de aantallen gymnologische tests die gedurende de vier periodes werden afgenomen.

period	tests		
	gymnastics	swimming	total number of gymnological tests
1. till autumn holidays	5	2	7
2. till Christmas holidays	5	2	7
3. till Easter holidays	7	2	9
4. till summer holidays	6	2	8

Tabel 4.5. : Distribution of the gymnastic and swimming tests over the *schoolyear*

De som van de scores van alle gymnologische tests werd beschouwd als het schoolvorderingencijfer van de leerling in de lichamelijke oefening over de betreffende periode.

Verwachting

De verwachting wordt uitgesproken dat als gevolg van de twee extra lessen lichamelijke oefening de experimentele groep in de loop van het schooljaar grotere vorderingen zal maken dan de controlegroep.

1. A.A.H.P.E.R., Youth Fitness Test Manual, revised edition, A.A.H.P.E.R., Washington, 1967.
2. Asmussen, E., Some physiological aspects of fitness for sport and work. Proc. Soc. Med. 62, 1969, 1160-1164.
3. Astrand, P.O., K. Rodahl, Textbook of Work Physiology, McGraw-Hill, New York, 1970.
4. Balke, B., A simple field test for the assessment of physical fitness, CARI Report, Oklahoma City; Civil Aeromedical Research Institute; Federal Aviation Agency, 1963.
5. Bass, B.H., Lung function tests. H.K. Lewis & Co. Ltd., London, 1962.
6. Biener, K., Zur Frage der Belastbarkeit von Schülern im Turnunterricht aus sportmedizinischer Sicht, Körpererziehung, 11, 1961, 63-69.
7. Bleeker,J., F.H. Bonjer, Een ijkapparaat voor het controleren en afregelen van ergometers, Polytechn. Tijdschrift A., 637, 1960.
8. Bonjer, F.H., Fysiologische methoden voor het vaststellen van belasting en belastbaarheid, Van Gorcum, Assen, 1965.
9. Bovend'eerdt, J., H.C.G. Kemper, K. Rijsdorp, J. Savelsbergh, A. Spijkers, H. Vlasblom, Onderzoek naar een physical fitness test, Thomas 6 en 7, 1973, 188-194, 211 en 222-229.
10. Brouha,L., The step test: A simple method of measuring physical fitness for muscular work in young men, Res. Quart., 14, 1943, 31-35.
11. Brozek, J., F. Grande, J.T. Anderson, A. Keys, Densiometric analysis of body composition: Revision of some quantative assumptions, Ann. N.Y. Acad. Sci., 110, 1963, 113.
12. C.A.H.P.E.R., Fitness Performance Test Manual, C.A.H.P.E.R., 1967, Ontario.
13. Cooper, K.H., A means of assessing maximal oxygen uptake, JAMA, 203, 1968, 201-204.
14. Cooper, K.H., Aerobics, Bantam Books, New York, 1969.
15. Cumming, G.R., R. Keynes, A fitness performance test for school children and its correlation with physical working capacity and maximal oxygen uptake. Can.Med. Ass.J., 96, 1967, 1262-1269.
16. Cumming, G.R., Correlation of physical performance with laboratory measures of fitness. In. Frontiers of fitness, ed. R.J. Shephard, C.C. Thomas, Springfield, 1971.
17. Cumming, G.R., Correlation of athletic performance and aerobic power

in 12 to 17 year old children with bone age, calf muscle, total body potassium, heart volume and two indices of anaerobic power, In: Pediatric Work Physiology, Proc. of the fourth Int. Symposium, ed. O. Bar-Or, Natanya,1973.

18. Doolittle, T.L., R. Bigbee, The twelve-minute run-walk: a test of cardio-respiratory fitness of adolescent boys. Res. Quart. 39, 1968, 491-495.
19. Drake, V., G. Jones, J.R. Brown, R.J. Shephard, Fitness performance tests and their relationship to the maximal oxygen uptake of adults, Can. Med. Ass. J., 99, 1968, 844-848.
20. Edholm, O.G., A growing edge in measurement of physical fitness, Proc. Roy. Soc. Med., 62, 1969.
21. Falls, H.B., A.H. Ismail, D.F. Macleod, Estimation of maximum oxygen uptake in adults from AAHPER Youth Fitness Test Items, Res. Quart., 37, 1966, 192-201.
22. Fleishman, E.A., The structure and measurement of physical fitness, Prentice-Hall, Englewood-Cliffs, N.Y., 1964.
23. Folia Civitatis, Gymnastiek, Amsterdam, 1 mei 1971, nr. 29.
24. Hansen, G. Vergleichende Untersuchungen über das Verhältnis der aeroben zur anaeroben Kapazität bei maximaler Ergometrischer Leistung, Schweizer. Z.f. Sportmedizin, 15, 1967, 68-75.
25. Hollmann, W., C. Bouchard, H. Venrath, G. Herkenrath, Vergleichende Untersuchungen über die Aussagekraft des Master-, Harvard-, Tuttle-, Schneider-, Hettinger- und Lian-Tests bezüglich der Kardio-pulmonalen Leistungsfähigkeit, Zschr. f. Kreislff. 54, 7, 1966, 647-657.
26. Kemper, H.C.G., A.J. Poulus, N. van der Helm, De invloed van een circuittraining in de lessen lichamelijke opvoeding op enkele morfologische en functionele kenmerken bij 12 en 13-jarige jongens, De Lich. Opv., 2, 1970, 47-57.
27. Kemper, H.C.G., P.J.G. Mettrop, A.N.P. Oud, L.W.C. Tavecchio, N. van der Helm, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor VWO en HAVO op lichamelijke en geestelijke groei en ontwikkeling, De Lich. Opv., 12 en 13, 1971, 543 en 604.
28. Kemper, H.C.G., Experimenten over de lichamelijke aspecten van de lichamelijke opvoeding bij 12 en 13-jarige jongens, Richting, 12, 1972, 236-243(a).

29. Kemper, H.C.G., Heart rate during bicycle ergometer exercise in satts per kilogram body weight of 12 and 13 year old boys, Z. Kinderheilk. 113, 1972, 161-169 (b).
30. Luyken, R., J.F. de Wijn, J.C.A. Zaat, Skeletbouw en lichaamsgewicht, een statistisch-somatometrisch onderzoek bij 8 tot 25-jarige jongens en mannen, Ned. T. Geneesk. 111, 10, 1967, 441-450.
31. Luyken, R., F.W.M. Luyken-Koning, G.B. Post, De invloed van een korte periode intensieve lichamelijke inspanning op de lichaamssamenstelling van adolescenten, Centraal Instituut voor Voedingsonderzoek, TNO, Zeist, Rapport nr. 2825, febr. 1969 (a).
32. Luyken, R., F.W.M. Luyken-Koning, De voedingstoestand van verdienend adolescenten in Rotterdam, Centraal Instituut voor Voedingsonderzoek, TNO, Zeist, Rapport nr. 2827, febr. 1969 (b).
33. Luyken, R., F.W.M. Luyken-Koning, De voedingstoestand van leerlingen van de Haven- en Vervoersscholen "Prof.Rutten", te Rotterdam, Centraal Instituut voor Voedingsonderzoek, TNO, Zeist, Rapport nr. 3023, dec. 1969 (c).
34. Luyken, R., F.W.M. Luyken-Koning, Somatometrische gegevens betreffende de voedingstoestand en lichaamssamenstelling van verdienende meisjes-adolescenten in W-Nederland en vergeleken met die van een referentiegroep, Centraal Instituut voor Voedingsonderzoek,TNO, Zeist, Rapport nr. 3137, maart 1970.
35. Maksud, M.G., K.D. Coutts, Application of the Cooper twelve-minute run-walk test to young males, Res. Quart. 42, 1, 1971, 54-59.
36. Malina, R.M., Exercise as an influence upon growth; Review and critique of current concepts, Clinical Pediatrics, vol. 8, no. 1, 1969, 16-26.
37. Master, A.M., The Master two-step test, Amer.Heart Jrnl.75, 1968, 809-837.
38. Mellerowicz, H., Der Kreislauf des Jugendlichen bei Arbeit und Sport, Basel, Karger, 1965.
39. Metz, K.F., J.F. Alexander, An investigation of the relationship between maximum aerobic work capacity and physical fitness in twelve to fifteen year-old-boys, Res. Quart. 41, 1970, 75-81.
40. Mocellin, R., J. Rutenfranz, Zur Frage der Linearität der Pulsfrequenzanstieges bei steigenden Leistungen, In: Proc.second Symp. of

Ped.Group of working physiology, University Prague, 1970, 15-20.

41. Mocellin, R., J. Rutenfranz, R. Singer, Zur Frage von Normwerten der körperlichen Leistungsfähigkeit (W_{170}) im Kindes- und Jugendalter, Z. Kinderheilk., 110, 1971, 140-165.
42. Novak, L.P., Age and sex differences in body density and creatinine excretion of highschool children, Ann. N.Y. Acad.Sci., 110, 1963, 545.
43. Olree, H., C. Stevens, T. Nelson, G. Agneville, R.T. Clarck, Evaluation of the AAHPER Youth Fitness Test. J. of Sports Med. and Phys. Fitness, 5, 1965, 67-71.
44. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor V.W.O. en H.A.V.O. op lichamelijke en geestelijke groei en ontwikkeling. Werkstuk Lab. voor Psychofysiologie, Univ. v. Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv. 59, 12 en 13, 1971, 543 en 604.
45. Parizkova, J., The impact of age, diet and exercise on man's body composition, Ann. N.Y. Acad. Sci., 110, 1963, 661.
46. Paschen, K., Tägliche Bewegungszeit in der Grundschule, Schorndorf, Verlag K. Hofmann, 1971.
47. Pieters, J.J.L., J.F. de Wijn, J.C.A. Zaat, Oriëntering over de relatie tussen vetheid en lichaamsmaten bij een groep getrainde sportlieden, Ned.T. Geneesk., 116, 53, 1972, 2384-2392.
48. Placheta, Z., A. Novak, V. Drazil, Z. Stich, Einige Probleme der Funktionstests bei submaximaler Belastung, Med. u. Sport, 5, 1971, 153-157.
49. Rarick, G.L., Exercise and growth, in: Science and medicine of exercise and sports, ed. W.R. Johnson, Harper & Row Publishers, New York, 1960.
50. Rasch, P.J., I.D. Wilson, The correlation of selected laboratory tests of physical fitness with military endurance. Milit. Med. 129, 1964, 256-258.
51. Rijsdorp, K., Cijfers en feiten inzake lichamelijke opvoeding, Wolters, Groningen, 1952.
52. Shephard, R.J., Endurance fitness. Univ. of Toronto, Toronto, 1968.
53. Shephard, R.J., C. Allen, A.J.S. Benade, C.T.M. Davies, P.E. di Prampero, R. Hedman, J.E. Merriman, K. Myhre, R. Simmons, The maximum oxygen intake. An international reference standard of

cardiorespiratory fitness. Bull. Wld.Hlth.Org., 38, 1968, 757-764.

54. Shephard, R.J., Learning, habituation and training, Int.Z.Angew. Physiol., 28, 1969, 38-48.

55. Simons, J., G. Beunen, M. Ostyn, R. Renson, P. Swalus, D. van Gerven, E. Willems, Constructie van een motorische testbatterij voor jongens van 12 tot 19 jaar door middel van factor-analysen. Sport, 49, 1, 1970, 3-21.

56. Tanner, J.M., Relationships of different bodily tissues during growth and in the adult, in: Diet and bodily constitution, Churchill Ltd. London, 1964.

57. Tanner, J.M., Growth and physique, in: IBP Handbook no. 9, Human Biology, a guide to field methods, ed. Weiner, J.S. and J.A. Lourie, Blackwell, Oxford, 1969.

58. Voorstel leerplan Rijksscholen voor lichamelijke oefening voor V.W.O., H.A.V.O., en M.A.V.O., Ministerie van Onderwijs en Wetenschappen, Den Haag, 1968.

59. Wahlund, H., Determination of the physical working capacity. Acta med. scand., suppl. 215, 1948.

60. Weiner, J.S., J.A. Lourie (eds), I.B.P., Handbook no. 9, Human Biology, a guide to field methods, Blackwell, Oxford, 1969.

61. Westerhof, R., docent didaktiek aan de A.L.O. te Groningen, 13-10-1972.

62. W.H.O., Exercise tests in relation to cardiovascular function, Techn. Report Series no. 388, Geneva, 1968.

63. Wiley, J.F., L.G. Shaver, Prediction of maximum oxygen intake from running performances of untrained young men, Res. Quart. 43, 1972, 89-93.

64. Willems, E.J., G. Beunen, M. Ostijn, R. Renson, J. Simons, P. Swalus, D. van Gerven, De differentiatie van krachtfactoren bij 12 tot 19-jarige jongens, Hermes, 6, 1971, 39-100.

65. Wilmore, J.H., A.R. Behnke, An antropometric estimation of body density and lean body weight in young men, J. Appl.Physiol., 27, 1, July 1969, 25-31.

4.2. PSYCHOLOGISCHE VARIABELEN

De doelstellingen uit het Voorstel leerplan Rijksscholen (1968) bieden aanknopingspunten, met behulp waarvan onderzocht kan worden welke invloed de lichamelijke oefening heeft op de geestelijke ontwikkeling. Tot nu toe is er weinig experimenteel onderzoek verricht om te toetsen of die doelstellingen realiseerbaar zijn en of de daarbij genoemde aspecten aan verandering onderhevig zijn, en zo ja, of deze effecten differentiëren als gevolg van extra lessen lichamelijke oefening. Bovendien blijken de resultaten uit onderzoekingen nogal tegenstrijdig te zijn, waardoor een minder duidelijk beeld is te vormen omtrent de effecten van de lichamelijke oefening op de geestelijke ontwikkeling (zie ook de inleiding).

Zo vonden Mulder (1962) en Hendrix (1968) onder andere een verbetering van het prestatievermogen en zelfvertrouwen als gevolg van een intensieve trainingsperiode op de Outward Bound School. In een onderzoek bij jongens van 12 tot 18 jaar (Tillman, 1965) werd gevonden dat de gemiddelde verandering van de experimentele groep (-1,81) wat betreft de schaal administratieve interesse van de Kuder Preference Record significant verschilde met de gemiddelde verandering van de controlegroep (+2,29). Dit betekende dat slechts 1 van de 28 gemeten psychologische variabelen een significant verschil vertoonde.

Wanneer bovenstaande doelstellingen als uitgangspunt worden genomen voor dit onderzoek, is het interessant om na te gaan wat de invloed is van een verschil in aantal lessen lichamelijke oefening, hetgeen wat betreft de geestelijke ontwikkeling mogelijk tot uitdrukking komt in een verandering van de prestatiebereidheid, zin tot samenwerken, zelfvertrouwen, doorzettingsvermogen, zelfkennis en zelf-acceptatie, en in een verandering van attitude ten aanzien van concepten waarop de lessen lichamelijke oefening van invloed zouden kunnen zijn.

4.2.1. S.A.G.S.

Het onderwijs in de lichamelijke oefening heeft ten doel bij te dragen tot de 'bevordering van de zin voor samenwerking' (Voorstel leerplan Rijksscholen, 1968).

Het Voorstel leerplan Rijksscholen doet vermoeden dat de zin voor

samenwerking aanwezig is en dat deze groter wordt door de lichamelijke oefening. Men zou daarom op grond van het Voorstel leerplan Rijksscholen mogen verwachten dat de zin voor samenwerking bij vijf lessen lichamelijke oefening in vergelijking met drie lessen lichamelijke oefening per week groter zal zijn.

Welke meettechnieken zijn beschikbaar om samenwerking te registreren?

1. observatie- en beoordelingsmethoden (Gage, 1971);
2. sociometrische methoden.

Bij observatie- en beoordelingsmethoden van groepen wordt aan de proefpersonen een complexe taak opgedragen, waarvan verwacht wordt dat de oplossing van de taak leidt tot een gemeenschappelijke uitvoering. Het verrichten van deze speciale 'veldtaken' kan vele uren in beslag nemen. Tijdens de uitvoering wordt het gedrag van de leden van de groep geobserveerd, beoordeeld en vastgelegd. Na afloop van de uitvoering wordt het gedrag geanalyseerd. Dit soort methoden wordt genoemd: schatting van de persoonlijkheid. Een bekend voorbeeld is het opzetten van een woon-industriegebied door een groep aspirant-officieren bij de officierskeuring van het Nederlandse leger. Het probleem bij het meten van het effect van lichamelijke oefening op de zin voor samenwerking is voor dit onderzoek enerzijds een praktisch probleem: hoe kan men betrouwbaar observeren in een groep leerlingen van ongeveer dertig personen, hoe kan men dit binnen het lesrooster realiseren, organisatieproblemen, enzovoorts? Anderzijds is het een methodologisch probleem: bestaat er een specifieke testsituatie waarin objectief en valide de zin voor samenwerking kan worden gemeten?

Het antwoord op beide vragen is niet voorhanden, daarom zullen we in het kort nagaan of de sociometrische methode een beter alternatief biedt.

De sociometrische methode

De sociometrie behandelt een belangrijk aspect van de persoonlijkheid, namelijk welke affectieve relatie een persoon heeft met anderen.

De sociometrische methode is een techniek welke dient om het interactiepatroon (= sociometrische structuur) binnen een groep op te sporen.

De sociometrische methode is door Moreno (1934, 1951, 1960) ontworpen. Aan de groepsleden wordt gevraagd wie als leider wordt beschouwd, met wie uit de groep ze het liefst met vakantie gaan, samenwerken, etc. en met wie beslist niet. Uiteraard kan men de vraagstelling kiezen afhankelijk van wat men in het onderzoek wil weten.

Door het optellen van scores komt men tot een populariteitsindex of afwijzingsindex. Men kan ook persoonlijkheidseigenschappen gebruiken bij dagelijks sociometrisch onderzoek. Dan stelt men bijvoorbeeld de vraag: 'wie vind je de intelligentste, en wie de domste van de klas'?

Door een analyse van de voorkeuren komt men tot een overzicht van de interactiepatronen, onderlinge verhoudingen en subgroepen binnen de groep. Men kan deze sociometrische structuur uitzetten in een diagram: het sociogram.

Hoewel de sociometrische methode in vergelijking met de observatie- en beoordelingsmethoden meer objectief scoorbaar is en bovendien snel en efficiënt is toe te passen, kleven er ook nadelen aan de sociometrische methode volgens Moreno.

1. Er is een beperkt oordeel. Niet alle leden van de groep hoeven een beoordeling te ontvangen;
2. het kiezen kan willekeurig geschieden, immers een keuze van personen moét gemaakt worden, ook als men liever geen voorkeur zou willen geven. Er kan niemand zijn met wie men het liefst op vakantie gaat;
3. een verschil in grootte van voorkeur (meer of minder voorkeur) komt niet tot uiting.

Omdat in een onderzoek naar de effecten van vijf lessen ten opzichte van drie lessen lichamelijke oefening per week waarschijnlijk sprake is van meer en minder, en om een waardering over alle kinderen te hebben, is gezocht naar een sociometrisch instrument dat bovengenoemde bezwaren niet kende.

Keuze van het instrument

De keuze viel op de Syracuse-Amsterdam-Groningen Sociometrische Schaal (Defares e.a., 1970), een vertaling van de Amerikaanse Syracuse Scales of Social Relations van Gardner en Thompson uit 1958.

De drie bezwaren zoals voornoemd bij de sociometrische methode van Moreno vervallen, immers:
álle personen ontvangen én geven een waardering, en daarnaast kan men de hoogte van de waardering variëren op een 5-puntsschaal. De S.A.G.S. kan gebruikt worden voor kinderen. Aan de kinderen wordt gevraagd om in de meest rechtse rij van vijf hokjes de naam te schrijven van de persoon met wie men het liefst van alle mensen wil praten als men erg verdrietig is, en in het meest linkse vakje de naam van degene met wie men in geen geval wil praten. Dit zijn de referentiepersonen. De leraar of de groepsleider leest vervolgens alle namen van de leerlingen uit de klas een voor een op, dit wil zeggen iedereen geeft aan, en ontvangt van de groepsleden één waardering.

Uit de gegevens kunnen onder meer de volgende scores berekend worden:

a. de afzonderlijke beoordelingen, dit levert per groepslid (n-1) scores op;
b. de individuele gemiddelden: de gemiddeld gegeven score: G
   de gemiddeld ontvangen score: O;
c. de individuele ratio-score: O : G = R;
d. het groepsgemiddelde van de beoordelingen, dit geeft aan hoe de groep als geheel over de groepsleden denkt: M;
e. vriendschapsstructuren, welke subgroepen kan men onderscheiden binnen de groep.

De ijking van de S.A.G.S. in Nederland (Defares e.a., 1970) werd uitgevoerd op circa 1200 leerlingen uit de zesde klas Lager Onderwijs.

Bij de ijking van de S.A.G.S. worden onderscheiden: G, het gemiddelde van de door één lid van de groep aan de anderen gegeven schaalwaarden; O, het gemiddelde van de door één lid van de groep van de andere leden van de groep ontvangen schaalwaarden; R, de verhoudingsscore, namelijk O : G.

De hertestbetrouwbaarheden van deze drie scores met een tijdsinterval van circa drie maanden bedragen respectievelijk: 0,89, 0,68 en 0,71. In het algemeen kan men stellen dat een hoge betrouwbaarheid wijst op een ongevoelig meetinstrument. Er kan namelijk moeilijk onderscheid worden gemaakt tussen geheugen en werkelijke verandering (Gage, '71).

Voor het onderzoek naar de effecten van vijf lessen lichamelijke oefening ten opzichte van drie lessen, en gezien het doel van het Voorstel leerplan Rijksscholen, is het wellicht interessant te onderzoeken of het groepsgemiddelde van de beoordelingen verandert. Uit het totaalgemiddelde van de groep kan men afleiden hoe de groepsleden globaal genomen over hun groep denken. Het is een indicatie voor de mate waarin de groep als bevredigend wordt gezien door de leden. Het is plausibel deze groepsscore in verband te brengen met begrippen als groepssfeer en groepscohesie (Defares e.a., 1970), begrippen die men uit de zin voor samenwerking uit het Voorstel leerplan Rijksscholen kan afleiden.

Verwachtingen

Zou lichamelijke oefening de zin voor samenwerking bevorderen en daarmee de groepssfeer en groepscohesie, dan mag men verwachten dat de groepsscore door lichamelijke oefening zal toenemen en een verschil in aantal lessen lichamelijke oefening eveneens tot uiting komt. Wij verwachten daarom dat het groepsgemiddelde zoals bepaald met de S.A.G.S. in de experimentele groep (de groep met vijf lessen lichamelijke oefening per week) zal stijgen ten opzichte van de controlegroep (de groep met drie lessen lichamelijke oefening per week).

4.2.2. NEUROTISCHE LABILITEIT

In het vooronderzoek (Oud e.a., 1971) is gebruikgemaakt van de Amsterdamse Biografische Vragenlijst (A.B.V.) ter bepaling van neurotische labiliteit en enige andere persoonlijkheidskenmerken (Wilde, 1963, 1970) om een maat te krijgen voor de aangepastheid van de proefpersonen. Ten aanzien van de overeenkomst met het begrip neurose zegt Wilde: 'Neurotische labiliteit is dat gemeenschappelijke persoonlijkheidskenmerk dat bepalend is voor de waarschijnlijkheid dat een individu een zodanig gedrag zal vertonen, dat hij onder één of meer van Scott's "research"-definities valt'. Dit betekent voor dit onderzoek dat neurotische labiliteit geacht wordt de kans te determineren dat de leerling valt onder de volgende als relevant beschouwde researchdefinities: ten eerste dat het individu blijkgeeft zich objectief ongelukkig en insufficiënt te voelen; ten tweede dat het individu faalt in de verwerkelijking van zijn eigen mogelijkheden.

Overeenkomstig het vooronderzoek (Oud e.a., 1971) luidt de veronderstelling, dat onder invloed van lichamelijke oefening de neurotische labiliteit zal afnemen en sterker zal afnemen bij een toename van het aantal lessen lichamelijke oefening. In het bovengenoemde onderzoek kon dit wat betreft de neurotische labiliteit niet worden aangetoond. De auteurs vermelden in de conclusies dat wellicht verschillen in sociaal-psychologische factoren tussen de experimentele en controlegroep een rol hebben gespeeld. De suggestie wordt gedaan om in een volgend onderzoek deze variantiebron te meten en wellicht een ander instrument te hanteren.

In aanmerking kwam de A.B.V. voor kinderen (A.B.V.-K), temeer daar dit instrument dat dezelfde schalen bevat als de 'volwassen' A.B.V., bedoeld is voor kinderen tussen 8 en 17 jaar oud (de leeftijd in dit onderzoek is gemiddeld 12,5 jaar) en misschien stabieler is dan de A.B.V. voor proefpersonen in de leeftijd van 13 tot 19 jaar (Wilde, 1963, 1970).

De correlaties tussen de instrumenten bleken echter laag te zijn, hetgeen twijfels opriep over de inhoud-validiteit van één van de twee instrumenten. De gebruikte aantallen proefpersonen waren echter laag. Omdat bovendien weinig bekend was over het gebruik (betrouwbaarheid, validiteit) van de pas uitgekomen A.B.V.-K (Wilde e.a., 1967) in experimentele situaties en daarmee over de mogelijkheid van vergelijking met andere onderzoekresultaten - hetgeen wel het geval is met de A.B.V. - werd besloten de A.B.V. als instrument ter bepaling van neurotische labiliteit ook in dit onderzoek te gebruiken.

Het instrument

Hoewel de A.B.V. in Nederland als genoegzaam bekend verondersteld mag worden, zal toch voor diegenen die het instrument niet kennen, het instrument in het kort beschreven worden.

De A.B.V. bevat vier schalen:

1. de N-schaal: neurotische labiliteit zoals gemanifesteerd in het hebben van psycho-neurotische klachten;
2. de NS-schaal: neurosomatische labiliteit zoals gemanifesteerd in het hebben van functionele (lichamelijke) klachten;
3. de E-schaal: sociale extraversie;

4. de T-schaal: de zelf-defensieve versus zelf-critische instelling.

De halveringsbetrouwbaarheid voor proefpersonen in de leeftijd van 13 tot 19 jaar was voor de N-schaal 0,90; de NS-schaal 0,39; de E-schaal 0,77; de T-schaal 0,84 . De testlabiliteit lag voor de N-, NS- en E-schaal boven de 0,90, terwijl deze voor de T-schaal 0,77 bedroeg.

Overigens wijst Wilde er wel op, dat bij herhaalde testafname de neiging bestaat tot het vertonen van 'minder onaangepaste' scores. Omdat de schalen, zoals hierboven gedefinieerd, enig misverstand kunnen opleveren, is het wellicht inzichtelijker, mede gezien de probleemstelling van dit onderzoek, de schalen in termen van sport te 'vertalen'. Nakken e.a. (1972) definiëren in hun zwemonderzoek de schalen als volgt:

1. N-schaal, hoge score: geen zelfvertrouwen, niet kunnen terugkomen bij een achterstand e.d.;
   lage score: zelfvertrouwen, kunnen terugvechten bij een achterstand, etc.;
2. NS-schaal, hoge score: grote aandacht voor lichamelijke wissewasjes,
   lage score: niet uit het veld te slaan door vermoeidheid, spierpijn, etc.;
3. E-schaal, hoge score: gemakkelijk omgaan met andere mensen,
   lage score: gesloten zijn, zich zelf isoleren;
4. T-schaal, hoge score: niet reëel inzicht hebben in eigen capaciteiten,
   lage score: realistisch inzicht hebben in eigen capaciteiten.

Verwachtingen

Met behulp van deze omschrijving en op grond van de doelstellingen en pedagogische intenties van de leraar lichamelijke oefening wordt verwacht dat de N-, NS- en T-scores zullen afnemen, en dat de E-score zal toenemen onder invloed van lessen lichamelijke oefening, alsmede dat deze veranderingen groter zullen zijn bij meer lessen lichamelijke oefening.

4.2.3. DE BEREIDHEID TOT HET LEVEREN VAN PRESTATIES

In de inleiding (zie 1.) is als een van de doelstelling-aspecten van

de lichamelijke oefening genoemd 'het vergroten van de bereidheid en de geschiktheid tot het leveren van prestaties'. Wij veronderstellen dat, vanuit deze doelstellingen gezien, met prestaties bedoeld worden onder meer die prestaties die liggen op het terrein van de lichamelijke activiteiten; dit betekent dat de bestudering van de geschiktheid tot het leveren van prestaties onder andere ligt op het gebied van het antropometrisch/morfologisch en fysiologisch gedeelte van het onderzoek. De belangstelling in dit deel gaat echter uit naar de bereidheid tot het leveren van prestaties, het motivatieprobleem.

In het vooronderzoek (Oud e.a., 1971) is als maat voor de prestatiemotivatie gebruikgemaakt van schoolvorderingencijfers. Ten aanzien van het schoolvorderingencijfer lichamelijke oefening werd wel een significant verschil gevonden tussen de experimentele groep en de controlegroep, maar niet ten aanzien van de schoolvorderingencijfers voor taal en rekenen. Het is echter de vraag of a) schoolvorderingencijfers wel de mogelijkheid bieden om (zelfs zeer geringe) verschillen in motivatieverandering weer te geven - een vraag naar de gevoeligheid van het meetinstrument; b) de schoolvorderingencijfers, gegeven door verschillende docenten, wel betrouwbaar genoeg kunnen worden geacht voor een vergelijking; c) verschillen in schoolvorderingencijfers inderdaad wel juiste indicaties zijn voor (het meten van) veranderingen in prestatiemotivatie - een validiteitsprobleem doet zich gelden. Daarom is besloten de prestatiemotivatie in dit onderzoek op een andere manier te meten en te zoeken naar een instrument dat meer adequaat aan onze vraagstelling beantwoordt. Drie mogelijkheden zijn voorhanden: 1) observatietechnieken; 2) projectietechnieken; 3) vragenlijst-methoden.

ad 1) Observatietechnieken zijn in dit experiment minder geschikt om gedragsveranderingen van de proefpersonen te meten. Vier klassen met elk gemiddeld 20 leerlingen zijn als observatie-object een te complex geheel om tot een objectieve, betrouwbare en efficiënte observatie van iedere leerling te komen. Bovendien bevatten de resultaten uit observaties bijna onvermijdelijk een subjectief element, zelfs al wordt er gewerkt met een beperkt aantal nauwkeurig omschreven gedragscategorieën.

'Niet het feit dat observatiemethoden geen inzicht kunnen

'verschaffen, want dat kunnen ze wel en soms veel dieper dan de tests, 'maar het gebrek aan objectiviteit en het daaruit voortvloeiende ge- 'brek aan betrouwbaarheid en validiteit zijn de oorzaak, dat zij als 'tests niet bruikbaar zijn' (De Zeeuw, 1971). Ook Drenth (1971) wijst op de geringe mate van objectiviteit van observaties, alhoewel hij stelt dat observatietests tot een verantwoord peil van objectiviteit gebracht kunnen worden door a) zoveel mogelijk aan te sluiten bij direct waarneembaar gedrag; b) met behulp van exact voorgeschreven verwerkingsprocedures de objectiviteit te bevorderen.

ad 2) Bij de projectiemethoden vindt men hetzelfde probleem: over het algemeen een geringe mate van betrouwbaarheid en objectiviteit (De Zeeuw, 1971; Drenth, 1971). Bovendien geldt dat voor zowel de projectiemethoden als voor de observatietechnieken veel opleidingstijd noodzakelijk is en ook de afname en de uitwerking aanzienlijk veel tijd in beslag nemen.

ad 3) Deze methode wordt gezien als een objectieve wijze van meten, waarbij de onderzochte door middel van geprecodeerde antwoorden een beschrijving geeft van zich zelf. Probleem hierbij is dat van de falsificaties: de onderzochte kan namelijk uit onwil of onvermogen een inadequate beschrijving van zich zelf geven. Ondanks vele pogingen is het eigenlijk nooit gelukt het falsificatieprobleem volledig onder controle te krijgen. Verder kan bij het beantwoorden van de items ook een vertekend beeld van de onderzochte ontstaan door een bepaalde respons 'set' of respons 'style', of door het geven van sociaal wenselijke antwoorden. Het falsificatieprobleem speelt vooral een rol bij de hantering van het inventarisatieprincipe, dat overigens niet zo bruikbaar bleek. Pas toen dit principe werd losgelaten en men uitging van de gedachte dat aan een consistente beantwoording van items bij bepaalde categorieën mensen een gemeenschappelijke karakteristiek ten grondslag ligt, werd het mogelijk vragenlijsten op een meer zinvolle wijze te hanteren (De Zeeuw,1971; Wilde, 1970). Het beantwoorden van de items zelf werd uitgangsgedrag, en niet meer de inhoud van de items. Hierbij bleven evenwel de respons 'set' en respons 'style' en het geven van sociaal wenselijke antwoorden een rol spelen. Toch bleken de vragenlijsten bruikbare instrumenten; ze hebben dan ook ruimschoots bewezen dat ze waarde hebben ondanks de bovengenoemde, niet geheel controleerbare invloeden.

Prestatiemotivatietest voor kinderen

Bovenstaande overwegingen hebben ons doen besluiten gebruik te maken van de vragenlijstmethode en onze keuze is daarbij gevallen op de prestatiemotivatietest voor kinderen (de PMT-K, Hermans, 1971).

In 'Prestatie-motief en faalangst in gezin en onderwijs' ontwikkelde Hermans (1971) een instrument (de PMT-K) ter meting van:

. het prestatiemotief: de tendens om het in taaksituaties goed te doen (P score);

. de negatieve faalangst: het disfunctioneren tijdens taaksituaties (F - score)

. de positieve faalangst: het beter functioneren tijdens taaksituaties (F+ score);

. de sociale wenselijkheid: de tendens om meningen, houding en denken prijs te geven voor groepsnormen (S.W. score).

Uitgangspunt is voor Hermans de 'social learning theory' zoals die door Rotter (1971 a en b) wordt gepresenteerd en waarin individuele verschillen en situatiekenmerken met elkaar in relatie worden gebracht. Afhankelijk van zijn 'reinforcement history' heeft het individu de verwachting dat een bepaald gedrag al of niet tot reinforcement zal leiden. Omdat bepaalde gedragspatronen kenmerkend zijn voor bepaalde typen mensen, betekent dit dat het mogelijk is tot classificatie van individuen te komen volgens bepaalde eigenschappen. En omdat individuen zich in eenzelfde situatie anders gedragen, kan volgens Hermans de aanname van bepaalde motieven verondersteld worden. Deze motieven bieden de mogelijkheid mensen met elkaar te vergelijken, waarbij in het model wordt uitgegaan van de interacties tussen classificaties van individuen en situaties. Gemotiveerd gedrag wordt dan gezien als: 'een functie van de interactie tussen de sterkte van een bepaald dispositioneel motief en de mate waarin het individu de verwachting heeft dat het doel van dat motief in een bepaalde situatie bevredigd zal worden (Hermans, 1971).

De prestatiemotivatietest voor kinderen (de PMT-K) bestaat uit 33 prestatiemotivatie items (P score), 15 negatieve en 17 positieve faalangst items (F- en F+ score) en 16 sociale wenselijkheids-items (S.W. score). De test is geschikt voor afname bij proefpersonen van 10 tot

16 jaar.

Als criteria ten aanzien van de validiteit werden gedrags- en belevingscriteria gehanteerd zoals rapportcijfers, het vroegtijdig staken van de studie, plezier in het studeren. Ten aanzien van rapportcijfers waren de validiteitscoëfficiënten wel significant maar niet consistent over de verschillende groepen; voor MAVO-leerlingen bijvoorbeeld minder hoog. Toch wijzen de resultaten duidelijk in één richting: hoog gemotiveerden hadden hogere rapportcijfers dan laag gemotiveerden.

Ook ten aanzien van de overige criteria bleken de PMT-K schalen voldoende validiteit te bezitten en een goed discriminerend vermogen te hebben.

De test-hertest correlatie is gemiddeld +0,64 over de 4 schalen, hetgeen in overeenstemming is met het standpunt dat motieven relatief stabiele persoonlijkheidsdisposities zijn, hetgeen voor dit onderzoek als minder gunstig kan worden gezien omdat wij juist in verandering zijn geïnteresseerd. Uit longitudinaal onderzoek blijkt dat voor kinderen het prestatiegedrag op intellectueel gebied stabiel is (Kagan e.a., 1962). Echter ten aanzien van prestaties op het gebied van de lichamelijke oefening (athletic achievement) is de stabiliteit uitgesproken negatief voor jongens en positief voor meisjes. De correlaties wat betreft deze lichamelijke oefening tussen de leeftijd 3-6 jaar en 6-10 jaar, en die tussen 6-10 jaar en 10-14 jaar bedroegen namelijk voor jongens respectievelijk -0,41 en -0,26 en voor meisjes respectievelijk +0,16 en +0,30. Kagan en Moss (1962) zeggen verder, dat onder de jongens van 10 tot 14 jaar weinigen waren, die zowel sport als school beiden als terrein voor prestatiemogelijkheid kozen.

De uiteindelijke schalen van de PMT-K zijn gebaseerd op een representatieve landelijke steekproef waarvan bijna vierduizend deelnemers voor de normerings-tabellen gebruikt konden worden.

Inhoudelijk is de prestatiemotivatietest voor kinderen vooral gericht op intellectueel gebied en hij sorteert daar het maximum aan effect. Het is niet duidelijk of de prestatiegemotiveerdheid op intellectueel gebied samenhangt met de prestatiegemotiveerdheid voor wat betreft de lichamelijke activiteiten en inspanningen. Dat wij

ondanks deze onduidelijkheid en ondanks het bezwaar dat motieven relatief stabiele persoonlijkheidsdisposities zijn, toch besloten hebben gebruik te maken van de PMT-K, vindt zijn oorzaak in het feit:

- dat er momenteel geen tests bestaan voor het meten van de prestatiemotivatie op het gebied van de lichamelijke activiteiten bij kinderen;
- dat er een mogelijke transfer bestaat van de geneigdheid tot het leveren van lichamelijke prestaties naar de geneigdheid tot het leveren van prestaties op intellectueel gebied;
- dat wij op grond van de doelstellingen van de lichamelijke oefening wel geïnteresseerd moeten zijn in eventuele veranderingen in de prestatiemotivatie.

Verwachtingen

Ten aanzien van de PMT-K zouden wij dan als mogelijk gevolg van lessen lichamelijke oefening kunnen verwachten dat een vergroting van de bereidheid tot het leveren van prestaties en het zich leren richten op het bereiken van een prestatie tot uitdrukking komt in een verhoging van de P score (prestatiemotivatie). Als lichamelijke activiteiten en oefening van invloed zouden zijn op het verkrijgen van bewegingszekerheid, zelfvertrouwen, het leren kennen van eigen kunnen en beperktheid, dan zou dat een mogelijke verhoging op de F+ score (positieve faalangst) en een mogelijke verlaging op de F- score (negatieve faalangst) tengevolge kunnen hebben. En als vijf lesuren lichamelijke oefening meer effect zouden sorteren dan drie lesuren, mag verwacht worden dat dit tot uitdrukking komt in het feit dat bij de groep met vijf lessen een grotere verhoging of verlaging op de PMT-K schalen valt waar te nemen dan bij de groep met drie lessen lichamelijke oefening. De S.W. score zal dienen als controleschaal, zoals deze ook door de schrijver is bedoeld.

4.2.4. ATTITUDEN TEN AANZIEN VAN CONCEPTEN WAAROP DE LICHAMELIJKE OEFENING INVLOED ZOU KUNNEN HEBBEN

In de inleiding vinden wij dat het onderwijs in de lichamelijke oefening ten doel heeft bij te dragen tot:

- bevordering van de zin tot samenwerken;
- kennismaking met waardevolle vormen van actieve recreatie;
- vergroting van de bereidheid en geschiktheid tot het leveren van prestaties.

Als nu op grond van lessen lichamelijke oefening veranderingen optreden, is het zinvol om tevens na te gaan in hoeverre deze veranderingen bij de proefpersonen gepaard gaan met wisselingen in attituden ten aanzien van concepten die met de hierboven genoemde aspecten verband houden. Immers, als de lichamelijke oefening bij voorbeeld de zin tot samenwerking bevordert, mag het als voor de hand liggend worden beschouwd dat de attitude ten aanzien van samenwerken tegelijkertijd aan verandering onderhevig is.

Semantische differentiaal

Omdat het gaat om de attitude ten aanzien van een aantal verschillende concepten waarop de lichamelijke oefening van invloed zou kunnen zijn, is het noodzakelijk een instrument te kiezen waarin een belangrijk deel van die concepten zijn opgenomen of kunnen worden opgenomen.

Onze keuze is daarbij gevallen op de semantische differentiaal van Osgood (Osgood e.a., 1957). Deze heeft namelijk het voordeel dat men zelf kan kiezen welke concepten men wil laten beoordelen op een aantal tegengestelde adjectief paren. Dit betekent voor ons dat een aantal begrippen kan worden gekozen, af te leiden uit de doelstellingen van de lichamelijke oefening.

In de 'Measurement of meaning' geven Osgood e.a. (1957) een methode aan voor het meten van de semantische betekenissen van woorden en concepten; daarbij verstaan zij onder semantische betekenis de relaties tussen tekens (signs) ten aanzien van hun significates (dat is iedere stimulus die in een gegeven situatie regelmatig en betrouwbaar een voorspelbaar gedragspatroon produceert (Osgood e.a., 1957)).

Door middel van scoring op een schaal met tegengestelde adjectieven kan worden aangegeven hoe iemand een bepaald object (concept) beoordeelt. Uit factoranalyse over de adjectief paren kwamen een aantal factoren te voorschijn; afhankelijk van de lading van de gekozen adjectieven op een bepaalde factor zijn de individuele schalen te

summeren tot factor-scores. De belangrijkste door de schrijvers genoemde factoren zijn:

- de evaluatiefactor met adjectief paren, zoals goed-slecht, waardevol-waardeloos, eerlijk-oneerlijk;
- de potentiefactor met schalen, zoals groot-klein, dik-dun, zwaar-licht;
- de activiteitsfactor met adjectief paren, zoals snel-langzaam, actief-passief, heet-koud.

Deze factoren bezaten respectievelijk ± 70, 15 en 12% van de gemeenschappelijke variantie. Wel blijkt uit ander onderzoek dat de factorladingen per concept kunnen verschillen, afhankelijk van welke groepen of klassen een oordeel uitspreken, zodat voorzichtigheid bij het gebruik geboden blijft (Denmark e.a., 1972; Heaps, 1972; Rosenbaum e.a., 1971).

Over de semantische differentiaal wordt gezegd (Osgood e.a.,1957) dat deze een nieuwe benadering en rationale geeft voor het meten van attituden, waaronder neigingen tot met name evaluatieve responsen worden verstaan, die betrekking hebben op toenadering of vermijding, of op een gunstig of ongunstig oordeel. Wanneer attituden op deze wijze worden opgevat, is het mogelijk een nulpunt of neutraal punt te veronderstellen, waardoor de mogelijkheid ontstaat tot registratie van zowel de richting als de intensiteit van de attituden (Krech e.a., 1962; Kenyon, 1968). Dit laatste is vooral van belang bij het meten van attitudeveranderingen.

Over de veranderlijkheid van attituden merken Krech e.a. (1962) op dat deze afhangt van:

- de kenmerken van reeds bestaande attituden;
- de persoonlijkheid van het individu;
- zijn verwantschap met de groep.

Wat betreft de veranderingen zelf zeggen zij dat deze veroorzaakt worden door:

1. kennismaking met de nieuwe informatie;
2. veranderingen van het individu wat betreft zijn verwantschap met de groep;
3. een gedwongen gedragswijziging ten opzichte van het object;
4. procedures die veranderingen in de persoonlijkheid teweegbrengen.

Gezien vanuit de factorstructuur van de semantische ruimte, waarbij de evaluatiefactor het grootste deel van de variantie voor zich opeist, lijkt het aannemelijk om attituden en attitudeveranderingen te registreren met behulp van de evaluatieve dimensie van die semantische ruimte. Hierbij dient opgemerkt te worden dat bepaalde concepten eerder een evaluatieve respons schijnen op te roepen dan andere, en dat de evaluatieve schalen minder stabiel en meer ontvankelijk zijn voor variatie over verschillende concepten dan de andere schalen (Komorita e.a., 1967).

De test-hertest-betrouwbaarheid van deze evaluatieve of attitudeschaal, gemeten met een interval van vijf weken, bedroeg gemiddeld +0,91 (range 0,88 tot 0,93). Ten aanzien van validiteit werden significante correlaties gevonden tussen de attitudescore van S.D. en een Thurstone scale (0,74 tot 0,91); ook de rangcorrelatie met een 14 item Guttman scale was significant (rho + 0,78) (Osgood e.a., 1957). Onderzoekingen gedaan bij kinderen tonen aan dat metingen van attitudes verricht met de S.D. betrouwbaar zijn, in ieder geval vanaf negen jaar, en dat bij het hoger worden van hun leeftijd ook de betrouwbaarheidscoëfficiënten groter worden (Di Vesta e.a., 1966; Maltz, 1963).

Keuze van de concepten

In het vooronderzoek (Oud e.a., 1971) werden twintig concepten gekozen die volgens een team, bestaande uit drie psychologen en een leraar lichamelijke oefening, belangrijk leken in verband met de doelstellingen van de lichamelijke opvoeding. Deze concepten waren als volgt verdeeld over vijf categorieën:

a. 7 concepten die verband hielden met de bereidheid en geschiktheid tot het leveren van prestaties;
b. 2 concepten die verband hielden met samenwerken;
c. 2 concepten die verband hielden met hygiëne;
d. 4 concepten die verband hielden met recreatie en vrijetijdsbesteding;
e. 5 concepten die verband hielden met leerprestaties.

Uit factoranalyse, berekend over de gegevens van de voortest uit dat vooronderzoek, kwamen vijf factoren te voorschijn; ze zijn als volgt benoemd en daarbij zijn telkens die concepten vermeld, welke

de hoogste ladingen op deze factoren vertoonden (zie tabel 4.6.):

1. leerprestaties met de concepten rekenen en lezen;
2. vrijetijdsbesteding met de concepten hobby's en sport;
3. samenwerken met de concepten elkaar helpen en samenwerken;
4. prestatiebereidheid met de concepten vroeg opstaan en doorzettingsvermogen;
5. prestatiegeschiktheid met de concepten transpireren en inspanning.

factors	concepts	factor loadings
samenwerking (cooperation)	elkaar helpen (helping each other	.77
	samenwerken (cooperating)	.77
prestatiegeschiktheid (performance-capability)	inspanning (exertion)	.79
	transpireren (perspiring)	.78
leerprestaties (mental performances)	rekenen (arithmatics)	.84
	leren (learning)	.80
prestatiebereidheid (performance willingness)	vroeg opstaan (getting up early)	.73
	doorzettingsvermogen (perseverance)	.67
recreatie in vrije tijd (recreation)	hobby's (hobbies)	.73
	sport (sports)	.70

Tabel 4.6. : the five factors and the two concepts with the highest loadings on each factor (total variance: 71%); the results come from the factor analysis of the data of the pilot study by Oud e.a. (1971) with 105 subjects.

Uit deze factoranalyse blijkt dat de categorie hygiëne in de factorstructuur van die twintig concepten niet voorkomt; de andere vier categorieën zijn wel duidelijk te herkennen, terwijl de categorie prestatiebereidheid en -geschiktheid zich in twee factoren opsplitst. Gezien de overeenkomst tussen de gevonden factorstructuur en de tevoren genoemde categorieën, en daarbij ook nog de mogelijkheid tot reductie van twintig naar tien concepten, is besloten om in dit onderzoek voor de vijf categorieën alleen de tien bovenstaande concepten te gaan gebruiken.

In dit onderzoek zal voor de meting van attitudeveranderingen ten aanzien van deze tien concepten gebruikgemaakt worden van de standaard semantische differentiaal (Jansen e.a., 1966). Deze interlandelijke standaard semantische differentiaal (S.S.D.) is eveneens in het vooronderzoek (Oud e.a., 1971) gebruikt, en bevat de volgende schalen:

voor de evaluatiefactor:

prettig	- - naar	gelukkig	- - ongelukkig
goed	- - slecht	aardig	- - onaardig

voor de potentiefactor:

groot	- - klein	zwaar	- - licht
sterk	- - zwak	hard	- - zacht

voor de activiteitsfactor:

actief	- - passief	boeiend	- - saai
vlug	- - langzaam	snel opgewonden	- - kalm

Voor een overzicht, zie bijlage IV.

Ten aanzien van attitudeveranderingen werden hierboven een viertal oorzaken genoemd. Gezien vanuit de leefwereld van leerlingen uit brugklassen zouden wij het volgende kunnen opmerken. De overgang van lagere school naar voortgezet onderwijs geeft een scala van nieuwe ervaringen. Het feit dat het in dit onderzoek nieuwe klassen betreft die zich nog als groep moeten ontwikkelen, kan een bijzonder licht werpen op de vorming en verandering van attituden. Van bijzonder belang is natuurlijk vooral het lesprogramma, waarin zowel drie als vijf uur gereserveerd is voor de lessen lichamelijke oefening voor de controle- respectievelijk de experimentele groep.

Verwachtingen

Verwacht wordt dan ook dat de extra lessen van invloed zullen zijn op de attituden betreffende concepten die betrekking hebben op de doelstellingen van de lichamelijke opvoeding. Voor de experimentele groep zou dat betekenen dat deze de genoemde concepten positiever gaat evalueren, zodanig, dat in vergelijking met de controlegroep de waardering voor die concepten bij de experimentele groep meer vooruitgaat. Voor de concepten hobby's en sport zou dit bovendien

een bevestiging zijn van de resultaten uit het vooronderzoek (Oud e.a., 1971).

4.2.5. SCHOOLATTITUDE

In het Voorstel leerplan Rijksscholen (1968) wordt als een van de doelstellingen van de lichamelijke oefening genoemd: 'vergroting van de bereidheid en geschiktheid tot het leveren van prestaties'. Naast de reeds genoemde operationalisaties van deze doelstelling door middel van de P-schaal van de PMT-K en attitudemetingen met behulp van de semantische differentiaal, kan men aannemen dat de eventuele effecten op de lichamelijke oefening eveneens zullen doorwerken in de attitude die de leerlingen vertonen ten aanzien van de school waarop zij zijn. Immers, de school is een instituut waar bereidheid en geschiktheid tot het leveren van (intellectuele) prestaties voortdurend een rol spelen. Op grond van het voorgaande is een attitudemeting ten aanzien van school als afhankelijke variabele binnen dit onderzoek opgenomen.

Ander onderzoek

Auria e.a. (1967) vermelden de resultaten van een factoranalyse op een satisfactieschaal bestaande uit 60 items, waarmee ze de tevredenheid van de leerling met de aan de school verbonden praktijken en activiteiten maten. Ze verkregen daarbij zeven onafhankelijke factoren. De vier best identificeerbare factoren dekten 72% van de variantie en hadden betrekking op:

. de schoolorganisatie;
. de aandacht welke de school besteedde aan sociale ontwikkeling;
. de inhoud van de leerstof en de normen;
. het gedrag van de onderwijzer.

In dit onderzoek zal mogelijkerwijs via manipulering van de leerstof de derde genoemde factor bij Auria e.a. (1967) een verandering op de mate van satisfactie bij de leerling tot stand komen.

Neal e.a. (1967) vonden in een onderzoek dat meisjes een positievere attitude dan jongens vertoonden ten opzichte van onderwijzer, schoolgebouw, klas, boeken lezen, schoolboeken in het algemeen en

het volgen van regels. Bovendien stelden Neal e.a. (1967) vast dat de attituden ten opzichte van schoolzaken negatiever werd naarmate men in een hogere klas zat. Flanders e.a. (1968) hebben op drie tijdstippen in een jaar bij 3.000 leerlingen de attitude jegens de onderwijzer gemeten, zijn eerlijkheid van belonen en straffen, zijn competentie en de attitude ten aanzien van 'schoolwork'. In de loop van het jaar bleek de attitude significant negatiever te worden. Tevens vonden zij dat het negatiever worden van de schoolattitude sterker wordt wanneer falen op school wordt ervaren als te zijn veroorzaakt door krachten buiten de eigen controle. Kuethe (1968) kwam tot de bevinding dat negatiever attitudes jegens de school, behalve door de houding thuis, vooral veroorzaakt worden door het feit dat de school 'faalervaring' betekent, hetgeen de motivatie en interesse zou verminderen. In ons onderzoek is onder andere de 'negatieve faalangst' van de leerlingen gemeten, zodat wij kunnen nagaan of onder invloed van de onafhankelijke variabele deze faalangst verandert, en eventueel daardoor de schoolattitude.

Jackson e.a. (1967) en Flanders e.a. (1968) vonden echter geen verband tussen schoolattitude en studieresultaten. Koopmans (1970) vindt het een logische veronderstelling dat een kind dat op school zijn behoeften en interessen heeft kunnen ontplooien, tevreden is met de school en met wat het daar doet, en dat het positief staat ten opzichte van onderwijzer en school. Koopmans (1970) waarschuwt echter dat 'tevredenheid' evenzeer kan betekenen dat het kind zich heeft aangepast (aan het systeem, aan de eisen), als dat gesproken kan worden van een tegemoetkoming aan de behoeften van het kind zelf. Koopmans geeft een uitstekend overzicht van onderzoek op dit terrein, waarin hij onder meer concludeert dat in geen enkel onderzoek de mate waarin de leerling zijn behoeften en interessen heeft kunnen ontplooien als criterium is gehanteerd.

Keuze van het instrument

Helaas is het zo dat alle bovengenoemde onderzoeken buiten Nederland zijn uitgevoerd, en dat men met instrumenten heeft gewerkt die buiten Nederland zijn gevalideerd, zodat conclusies uit deze reeks onderzoeken niet zonder meer toepasbaar mogen worden geacht op de

Nederlandse situatie. Op grond van deze overwegingen is dan ook gekozen voor een door de Groot e.a. (1968) ontwikkeld meetinstrument ter bepaling van schoolattitude, uitgevoerd en geijkt in Nederland. Deze attitudemeting beperkt zich tot slechts enkele aspecten van 'school', namelijk de attitude ten aanzien van school als onderwijsinstelling. Allerlei andere sociale en socialiserende aspecten van het schoolleven, ten aanzien waarvan de leerlingen zeker attituden ontwikkeld zullen hebben, zijn buiten beschouwing gelaten. Het door de Groot e.a. (1968) ontwikkelde instrument biedt het voordeel dat de beweringen waaruit de attitudemeting is opgebouwd, zijn samengesteld op grond van interviews met leerlingen, zodat de kans dat de bewegingen de leerling zullen aanspreken, vergroot is. Bovendien bestond de groep waarop het instrument werd geijkt, uit 12 en 13-jarige leerlingen van V.W.O. scholen, een populatie die veel overeenkomst vertoont met de door ons onderzochte groepen. Deze attitudetest bestaat uit 22 evaluatieve beweringen of uitspraken over het naar school gaan in het algemeen. Van elke bewering of uitspraak moet aangegeven worden of men deze 'waar' of 'onwaar' vindt. Van de 22 items zijn er slechts tien van belang voor de scoring. De overige twaalf vormen zogenaamde 'buffer' items, gehandhaafd om de context van de wél meetellende items te laten bestaan. Enkele voorbeelden:

'ik zal blij zijn als ik van school af ben' waar niet waar;
'ik wou dat ik op een school was waar ze geen huiswerk gaven'
waar niet waar.

Rutgers van der Loeff e.a. (1967) vermelden dat het instrument een betrouwbaarheidscoëfficiënt heeft van 0,69. Dit is niet zo laag, als men bedenkt dat het gaat om slechts tien items.

Verwachting

Naar aanleiding van het voorgaande verwachten wij dat de leerlingen van de experimentele groep ten aanzien van de school een meer positieve attitude zullen ontwikkelen dan de leerlingen uit de controlegroep.

4.2.6. GYMNASTIEK, VOORKEUR EN RANGNUMMER

Werd onder schoolattitude de houding van de leerling ten opzichte

van school in het algemeen behandeld, in dit hoofdstuk gaat het meer om de waardering van de leerling voor bepaalde schoolvakken, in dit geval de les lichamelijke oefening.

Een van de doelstellingen van de lichamelijke oefening is het kennismaken met waardevolle vormen van actieve recreatie. Deze kennismaking is echter alleen dan zinvol wanneer het ook werkelijk bijdraagt tot een actievere vrijetijdsbesteding.

Ander onderzoek

Bij een vergelijking van een groep zeer actieve met een groep inactieve jongens in de leeftijd van 12 tot 16 jaar (Ruffer, 1965), bleek onder meer de attitude ten aanzien van de lichamelijke oefening in het schoolprogramma significant te verschillen. De attitude van de zeer actieve leerlingen was veel positiever dan die van de inactieve leerlingen.

Keogh (1963) vergeleek studenten met een zeer positieve en studenten met een zeer negatieve attitude ten aanzien van lichamelijke oefening. Deze attitude bleek parallel te lopen met de waardering voor de lichamelijke oefening die men op de 'highschool' had gekregen. Ook de beoordeling van de eigen lichamelijke vaardigheid ten opzichte van de leeftijds- en seksegenoten vertoonde een nauwe relatie met de attitude.

Opvallend hierbij was dat desondanks de mate van actieve recreatie niet duidelijk verschilde. De verklaring voor het ontstaan van een positieve of negatieve attitude zoekt Keogh (1963) in emotionele factoren die opgewekt worden door ervaringen in de gymnastiekles en die variëren van een positieve beoordeling door leeftijdsgenoten bij een geslaagde uitvoering van een oefening, tot negatieve gevoelens van onbekwaamheid door gebrek aan kracht.

Wij veronderstellen dat juist deze ervaringen in positieve zin zullen veranderen onder invloed van extra lessen, gezien het feit dat de prestaties in de gymnastiekles toenemen door meer lessen lichamelijke oefening, zoals in ons vooronderzoek werd aangetoond (Oud e.a., 1971).

Andere gronden voor onze veronderstelling dat de attitude ten aanzien van het vak lichamelijke oefening in positieve zin zal

veranderen onder invloed van extra lessen, kunnen gevonden worden in een onderzoek naar de attitude van beginnende eerstejaarsstudenten (Brumbach, 1965), bij welk onderzoek een tendens viel waar te nemen tot een positievere attitude ten aanzien van lichamelijke oefening naarmate die studenten gedurende langere tijd op de 'highschool' lichamelijke oefening hadden gehad.

Vooronderzoek

In het vooronderzoek (Oud e.a., 1971) werden alle concepten ten aanzien van recreatie en vrijetijdsbesteding door de experimentele groep na een jaar lang twee extra lessen lichamelijke oefening per week, significant hoger gewaardeerd dan door de controlegroep met de normale drie lesuren. Daarnaast bleek uit de resultaten van een hiërarchische clusteranalyse die op de attituden werd uitgevoerd, dat het in aanvang bij beide groepen aanwezige cluster 'actieve vrijetijdsbesteding' zich alleen bij de experimentele groep had gehandhaafd, en dat de controlegroep gymnastiek evalueerde als een der schoolvakken. Met andere woorden, de attitude ten aanzien van gymnastiek kwam bij de experimentele groep meer te liggen in de sfeer van de actieve vrijetijdsbesteding, hegeen als een zeer belangrijke verschuiving kan worden beschouwd.

Het instrument

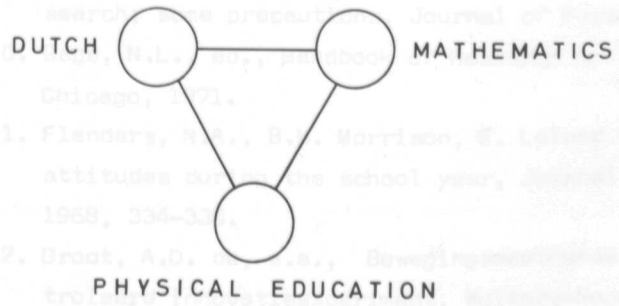
Vooral in de Verenigde Staten hebben een aantal onderzoekers zich beziggehouden met de constructie en toepassing van instrumenten ter meting van de attitude ten aanzien van lichamelijke oefening en sport (Wear, 1951, 1955; Keogh, 1962, 1963; Adams, 1963; Kenyon, 1968). Het feit echter dat deze instrumenten in het buitenland zijn ontwikkeld en geijkt, en over het algemeen zijn toegepast bij studenten, moet tot de conclusie leiden dat de instrumenten zelf, en de onderzoeksresultaten ermee behaald, in de meeste gevallen niet van toepassing mogen worden geacht op onze populatie van 12 en 13-jarigen. Tot op heden is geen van deze instrumenten in het Nederlands vertaald, laat staan geijkt of toegepast. In Nederland is op dit gebied nog nauwelijks iets ontwikkeld.

Binnen het kader van dit onderzoek is het vanzelfsprekend een

belangrijke zaak om te weten of de positie van het vak lichamelijke oefening aan het einde van de onderzoeksperiode bij de experimentele groep veranderd is in vergelijking met de positie van dit vak bij de controlegroep.

De Groot e.a. (1968) doen onder meer verslag van een methode om de relatieve voorkeur voor het vak meetkunde ten opzichte van een aantal andere schoolvakken te bepalen. Gebruikmakend van deze beschrijving hebben wij op twee manieren getracht de relatieve voorkeur voor het vak lichamelijke oefening te bepalen:

1. men laat de leerlingen op een vaste lijst van schoolvakken (alle vakken die ze hebben) een rangorde van voorkeur aangeven; het rangnummer van gymnastiek is dan de score.
2. met behulp van het triademodel van Coombs (1956) worden de vakken telkens in groepen van drie gepresenteerd, waarvan één vak als het meest prettige, en een ander als het minst prettige moet worden aangegeven (zie schema 4.7.). Door de herhaalde presentatie met twee andere vakken is het nu mogelijk na te gaan in hoeverre de leerling consequent is geweest, zodat men een eenvoudige betrouwbaarheidsmaatstaf heeft. De volgende vakken werden in deze voorkeurtest opgenomen: gymnastiek, wiskunde, geschiedenis, tekenen, Frans, aardrijkskunde en Nederlands. De betrouwbaarheid van deze methode bedraagt .88 (Rutgers van der Loeff e.a., 1967). De resulterende rangplaats van gymnastiek is hier de score.



Schema 4.7. : Coombs' triade model.

Verwachting

Hoewel de hoeveelheid onderzoekresultaten waaraan wij kunnen refereren niet erg groot is, verwachten wij dat de positie van het vak lichamelijke oefening bij de experimentele groep gunstiger zal zijn dan bij de controlegroep.

Literatuur

1. Adams, R.S., Two scales for measuring attitude toward physical education, Res. Quart. 34, 1, 1963, 91-94.
2. Auria, C., R.G. Frankiewicz, Components of school satisfaction among a group of ninth and tenth graders, in: Proceedings of the 75th Annual Convention of the A.P.A., 2, 1967, 321-322.
3. Brumbach, W.B., J.A. Cross, Attitudes toward physical education of male students entering the University of Oregon, Res. Quart. 36, 1, 1965, 10-16.
4. Coombs, C.H., Theory and methods of social measurement in: Group Dynamics, Research and Theory, ed. Cartwright, D.A. Zander, New York,1956, 471-535.
5. Defares, P.B., G.N. Kema, E. van Praag, J.J. van der Werff, Syracuse-Amsterdam-Groningen Sociometrische Schaal, uitgave voor research-doeleinden,1970.
6. Denmark, F.L., E.J. Shirk, R.T. Riley, The effect of ethnic and social class variables on semantic differential performance, Journal of social psychology, volume 86, 1972, 3-9.
7. Di Vesta, Fr.J., W. Dick, The test-retest reliability of children's ratings on the semantic differential, Educational and psychological measurement, 26, 1966, 605-616.
8. Drenth, P.J.D., De Psychologische test. Psychologische Monografieën, Van Loghum Slaterus, Deventer, 1971.
9. Heaps, R.A., Use of the semantic differential technique in research; some precautions, Journal of Psychology, 80, 1972, 121-125.
10. Gage, N.L., ed., Handbook of Research on Teaching, Rand McNally Co., Chicago, 1971.
11. Flanders, N.A., B.M. Morrison, E. Leland Brode, Changes in pupil attitudes during the school year, Journal of Ed. Psychology, 50, 1968, 334-338.
12. Groot, A.D. de, e.a., Bewegingsmeetkunde, Verslag van een gecontroleerd innovatieexperiment, Wolters-Noordhoff n.v., Groningen, 1968.
13. Hendrix, J.M.G., De Outward Boundschool; een exploratief onderzoek, doctoraalscriptie, Universiteit van Nijmegen, 1968.
14. Hermans, H.J.M., Prestatie motief en faalangst in gezin en onderwijs, Tevens handleiding bij de Prestatie motivatie test voor

kinderen (PMT-K), Swetz en Zeitlinger, Amsterdam, 1971.

15. Jackson, P.W., H.M. Lahaderne, Scholastic success and attitude toward school in a population of 6th graders, Journal of Ed. Psychology, 58, 1967, 15-18.
16. Jansen, J.M., A.J. Smolenaars, Kort verslag inzake een interlandelijk gestandaardiseerde semantische differentiaal, Ned. Tijdsch. v.d. Psychologie, 11, no. 3, 1966, 211-216.
17. Kagan, J., Moss, H.A., Birth to maturity, New York, Wiley, 1962
18. Kenyon, G.S. Six scales for assessing attitude toward physical activity, Res. Quart. 39, 3, 1968, 566-574.
19. Keogh, J., Analysis of general attitudes toward physical education, Res. Quart., 33, 2, 1962, 239-244.
20. Keogh, J., Extreme attitudes toward physical education, Res. Quart., 34, 1, 1963, 27-33.
21. Komorita S.S., R. Bass, Attitude differentiation and evaluative scales of the semantix differential, Journal of Personality and social Psychology, 6, 2, 1967, 241-244.
22. Koopmans, J.J.G., Motivatie van de leerling en attitude van de onderwijzer in het kader van het aptitude-treatment-interaction model, Doctoraalscriptie, Universiteit van Nijmegen, 1970.
23. Krech, D., R.S. Crutchfield, E.L. Ballachey, Individual in society, International student edition, Mc.Graw Hill Book Comp.Inc., 1962.
24. Kuethe, J.L., The teaching learning process, Glenview III.Scott., Foresman and company, 1968.
25. Maltz, H.E., Ontogenetic change in the meaning of concepts as measured by the semantic differential, Child development, 34, 1963, 667-674.
26. Moreno, J.L., Who shall survive? Washington D.C.: Nervous and Mental Disease Publication Co., 1934.
27. Moreno, J.L., Sociometry, experimental method and the Science of Society, New York, Beacon House, 1951.
28. Moreno, J.L., ed., The Sociometry Reader, Glencoe, Illinois: Free Press, 1960.
29. Mulder, H.F., Karaktervorming door Lichamelijke Inspanning, Paedagogische studieën, 39, 12, 1962.
30. Nakken, H., H. van der Spek, Zwemonderzoek, rapport betreffende

enkele psychologische en fysiologische aspecten bij wedstrijdzwemmers, Universiteit van Amsterdam, kring A'dam KNZB, 1969.

31. Neal, D.C., J.M.Proshek, School related attitudes of culturally disadvantaged school children, Journal of Ed. Psychology, 58, 1967, 238-244.
32. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor V.W.O. en H.A.V.O. op lichamelijke en geestelijke groei en ontwikkeling, Doctoraal werkstuk, Univers. v. Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opvoeding, 59, 12 en 13, 1971, 543 en 604.
33. Osgood, C.E., G.J. Suci, P.H. Tannenbaum, The measurement of meaning, University of Illinois Press, Chicago, 1957.
34. Rosenbaum, L.L., W.B. Rosenbaum, E.T. Mc Ginnies, Semantic differential factor structure stability across subject, concept, and time differences, Multivariate Behavioral Research,6, 4, 1971, 451-469.
35. Rotter, J.B., Social learning and clinical psychology, in: Hermans H.J.M. Prestatie motief en faalangst in gezin en onderwijs, Swetz en Zeitlinger, Amsterdam, 1971 (a).
36. Rotter, J.B., Some implications of social learning theory for the prediction of goal-directed behavior from testing procedures, in: Hermans H.J.M. Prestatie motief en faalangst in gezin en onderwijs, Swetz en Zeitlinger, Amsterdam, 1971 (b).
37. Rüffer, W.A., A study of extreme physical activity groups of young men, Res. Quart. 36, 2, 1965, 183-196.
38. Rutgers van der Loeff, P., E. Wesdorp-Uytenbogaart, School-attituden van HAVO-leerlingen, doctoraalwerkstuk, Universiteit van Amsterdam, 1967.
39. Snel, J., L.W.C. Tavecchio, Een onderzoek naar de relatie tussen de A.B.V.-K en de A.B.V., Intern Rapport, Laboratorium voor Psychofysiologie, Amsterdam, 1972.
40. Tillman, K., Relationship between physical fitness and selected Personality traits, Research Quarterly, 36, 4, 1965, 483-489.
41. Voorstel leerplan Rijksscholen, voor lichamelijke oefening, voor V.W.O., H.A.V.O. en M.A.V.O., Ministerie van Onderwijs en Wetenschappen, Den Haag, 1968.

42. Wear, C.L., The evaluation of attitude toward physical education as an activity course, Res. Quart. 22, 1951, 114-126.
43. Wear, C.L., Construction of equivalent forms of an attitude scale, Res. Quart. 26, 1955, 113-119.
44. Wilde, G.J.S., Neurotische labiliteit, gemeten volgens de vragenlijstmethode, F. van Rossen, Amsterdam, 1963.
45. Wilde, G.J.S., Neurotische labiliteit, gemeten volgens de vragenlijstmethode, 2e druk, F. van Rossen, Amsterdam, 1970.
46. Wilde, G.J.S., H. van Dijl, Voorlopige Handleiding bij de A.B.V. voor kinderen, F. van Rossen, Amsterdam, 1967.
47. Zeeuw, J. de, Algemene psychodiagnostiek 1, testmethoden, Swetz en Zeitlinger, Amsterdam, 1971.

hoofdstuk 5

STORENDE VARIABELEN

5.1. BIOLOGISCHE LEEFTIJD

- groei en ontwikkeling
- het begrip biologische leeftijd
- skeletleeftijd
- de methode van Tanner - Whitehouse - Healy
- meetprocedure
- verschillen in trainingseffect
- resultaten
- verwachting

LITERATUUR

5.2. BUITENSCHOOLSE LICHAMELIJKE ACTIVITEIT

- meetvoorwaarden en toepassingsmogelijkheden
- methoden voor het meten van lichamelijke activiteit

. meetprocedure stappentellers
. meetprocedure vragenlijst
. verwachting
. resultaten

LITERATUUR

5.3. LESGEEFGEDRAG VAN DE LERAREN LICHAMELIJKE OEFENING

. cognitieve en emotionele aspecten
. het vastleggen van de lessen op videoband

5.3.1. VAKTECHNISCHE BEOORDELING VAN HET LESGEEFGEDRAG

. keuze van de beoordelaars
. keuze van het systeem
. beoordelingsprocedure
. keuze van de lessen
. resultaten

5.3.2. SOCIAAL-EMOTIONELE BEOORDELING VAN HET LESGEEFGEDRAG

. raakpunten met ander onderzoek
. interactie binnen de schoolklas
. de tien categorieën uit het Flandersysteem
. ontwikkeling van een instrument en een methode van beoordeling
. keuze van de lessen
. keuze van de beoordelaars
. beoordelingsprocedure
. beschrijving van de matrix
. interpretatie van de matrix
. resultaten

5.3.3. ACTIVITEIT TIJDENS DE LESSEN LICHAMELIJKE OEFENING

. meetprocedure

. resultaten

5.3.4. LEERLINGBEOORDELING VAN HET LERAARGEDRAG

. het instrument

. resultaten

LITERATUUR

5. STORENDE VARIABELEN

De onafhankelijke variabele wordt in dit onderzoek gevormd door de twee extra lessen lichamelijke oefening in schoolverband per week.

In het algemeen wordt verondersteld dat de lessen lichamelijke oefening de lichamelijke en geestelijke groei en ontwikkeling, en het lichamelijke prestatievermogen van het kind gunstig beïnvloeden (Voorstel leerplan Rijksscholen, 1968). In dit verband mag van vijf lessen lichamelijke oefening per week dan ook een groter effect verwacht worden dan van de gebruikelijke drie lessen. Voor de leerlingen die tot de experimentele groep behoren, zullen deze extra lessen een verhoging van het lichamelijk activiteitsniveau binnen het lesrooster van de school met tweederde betekenen.

Alhoewel een dergelijke verhoging toereikend lijkt voor het aantonen van eventuele effecten van lichamelijke oefening, zoals omschreven in het leerplan, doen zich echter enige aspecten voor welke de eventuele effecten van extra lessen kunnen maskeren, en die daarom 'storende' variabelen worden genoemd.

Als storende variabelen beschouwen wij:

a. de biologische leeftijd (zie 5.1.);
b. de buitenschoolse lichamelijke activiteit (zie 5.2.);
c. het lesgeefgedrag (zie 5.3.).

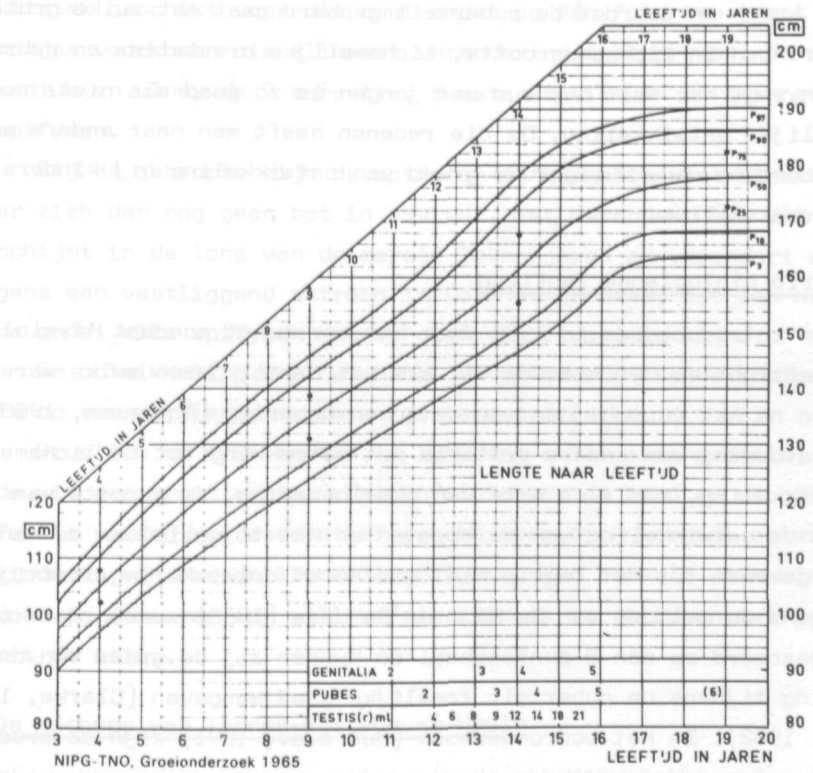
5.1. BIOLOGISCHE LEEFTIJD

Groei en ontwikkeling

Op het belang om bij kinderen de biologische leeftijd te bepalen, hebben Bouchard e.a. (1968), Clarke (1971, Cumming (1973) en Szabo (1972) reeds gewezen. Zij toonden aan dat tussen lichaamsbouw en lichamelijke prestaties enerzijds en biologische leeftijd anderzijds een grotere samenhang bestond dan met kalenderleeftijd.

Groei en ontwikkeling vinden voor elk kind niet in hetzelfde tempo plaats. De lichamelijke ontwikkeling geschiedt bij meisjes sneller dan bij jongens; meisjes zijn in de puberteit wat hun ontwikkeling betreft ongeveer twee jaar voor op jongens. Verder heeft geslachtelijke rijping een sterke samenhang met bepaalde groeifenomenen zoals de groeispurt (Van Wieringen e.a., 1968).

De snelheid van de lichamelijke ontwikkeling van kinderen met eenzelfde kalenderleeftijd (chronologische leeftijd) kan zeer ver uiteenlopen. Deze biologische variatie bij een gegeven kalenderleeftijd is klein op zeer jeugdige leeftijd, maar neemt toe tot een maximum wordt bereikt rond de puberteit. Daarna neemt de variatie weer af.



Figuur 5.1. : Height for chronological age relationship of Dutch boys from the N.I.P.G. growth study (Van Wieringen e.a., 1965).

Figuur 5.1. geeft naast de curve voor gewicht naar lengte uit het NIPG (TNO) groeionderzoek tevens de curve voor lengte naar kalenderleeftijd (Van Wieringen e.a., 1965). Waarden tussen het negentigste (P_{90}) en het tiende (P_{10}) percentiel kunnen als normale waarden worden

beschouwd voor de betreffende leeftijd. Uit de grafiek blijkt dat de kalenderleeftijd de lichamelijke ontwikkeling of rijping niet goed kan voorspellen omdat een jongen van bij voorbeeld dertien jaar zowel een lengte van 146 cm als van 166 cm mag hebben. Lengte blijkt ook geen nauwkeurige indicatie voor lichamelijke ontwikkeling te zijn, want een jongen van 155 cm lang kan zowel 14,8 jaar als 11,5 jaar oud zijn. Door de grote verschillen in leeftijd waarop kinderen in de puberteit komen en doordat de puberteit gepaard gaat met zulke grote veranderingen in lichaamsgrootte, lichamelijke prestaties en gedragpatronen,zegt het feit hoe oud een jongen is zo goed als niets over zijn lichamelijke ontwikkeling. Om die redenen heeft men naar andere methoden gezocht om nauwkeuriger de graad van ontwikkeling bij kinderen te kunnen vaststellen.

Het begrip biologische leeftijd

Crampton introduceerde in 1908 voor het eerst het concept 'fysiologische leeftijd' en definieerde dit als het aantal jaren welke waren verlopen na het verschijnen van de schaambeharing (Andersen, 1968). Door toevoeging van andere criteria met betrekking tot de lichamelijke ontwikkeling, zoals de mate van okselbeharing, de grootte van de testes, de ontwikkeling van de borsten en het tijdstip van de menarche, is men gekomen tot het begrip biologische of ontwikkelingsleeftijd.

Volgens de methode van De Wijn en De Haas (1960) worden deze criteria gescoord op een 5-puntsschaal en kunnen zij de graad van ontwikkeling tijdens de puberteit tamelijke goed aangeven (Clarke, 1971; Tanner, 1962). In het vooronderzoek (Oud e.a., 1971) zijn de proefgroepen volgens deze methode gescoord.

Er is echter behoefte aan een methode waarmee de lichamelijke ontwikkeling zowel voor, tijdens, als na de puberteit kan worden bepaald.

Skeletleeftijd

De ontwikkeling van het skelet kan worden gebruikt voor de bepaling van de biologische leeftijd. De skeletleeftijd geeft aan het stadium waarin de ontwikkeling van het skelet zich bevindt.

Elk botstuk doorloopt verschillende stadia van verandering van grootte, vorm en botvorming. Dergelijke veranderingen zijn te bestuderen op een röntgenfoto. De absorptie van de röntgenstraling is

afhankelijk van de dikte en samenstelling van de bestraalde materie; omdat botweefsel een grotere absorptie heeft dan kraakbeen, kunnen op de röntgenfoto de beenachtige centra onderscheiden worden van de kraakbeenachtige centra. De skeletleeftijd wordt vervolgens geschat uit het aantal botachtige centra en het ontwikkelingsstatium van elk.

In theorie kunnen alle delen van het skelet gebruikt worden om de skeletleeftijd vast te stellen. Het skelet van hand en pols is bij uitstek geschikt voor dit doel omdat op één foto een groot aantal botjes voorkomen die allemaal veranderingen ondergaan: de epifysen van radius en ulna, de 7 carpalia, 5 metacarpalia en 14 falangen met epifysen.

Bij de geboorte zijn de epifysen en carpalia niet zichtbaar omdat er zich dan nog geen bot in bevindt. Het merendeel van dit bot verschijnt in de loop van de eerste levensjaren en verandert daarna volgens een vastliggend patroon in vorm en grootte, tot aan het eind van de groei alle epifysen met de diafysen vergroeid zijn.

De ontwikkeling van het hand- en polsskelet op de röntgenfoto is op verschillende manieren gestandaardiseerd.

Gruelich en Pyle hebben in 1950 een atlas samengesteld, bestaande uit een groot aantal röntgenfoto's van de hand en de pols. Deze foto's worden representatief geacht voor zich normaal ontwikkelende kinderen van verschillende leeftijden. Men bepaalt de skeletleeftijd door de foto van een willekeurig kind te vergelijken met de foto's uit de atlas.

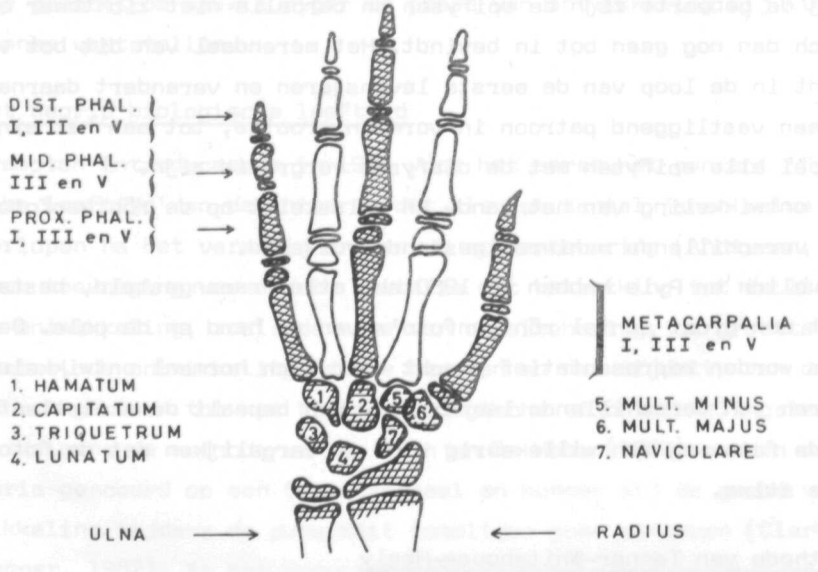
De methode van Tanner-Whitehouse-Healy

Wij hebben niet van een atlasmethode gebruikgemaakt, maar van een zogenaamde 'bone-specific or inspectional' methode (Malina, 1971) volgens Tanner-Whitehouse-Healy (1959, 1962). Daarbij wordt de skeletleeftijd bepaald uit in totaal twintig botjes: te weten radius en ulna (2), de metacarpalia 1, 3 en 5 (3), de proximale, mediale en distale falangen van 1, 3 en 5 (8), en alle carpalia (7) (zie figuur 5.2.).

Elk botstuk wordt onderscheiden in acht stadia, met uitzondering van de radius die negen stadia kent. Het stadium wordt bepaald door het betreffende botstuk te vergelijken met de beschrijving en

eventuele tekening. Van elk botstuk wordt de waarde van het begin- tot het eindstadium gesteld tussen 0 en 100.

Op biologische gronden veronderstelden Tanner e.a. (1959) dat pijpbeenderen en platte beenderen door verschillende factoren werden beinvloed. Een eenvoudig gemiddelde uit alle botjes zou een overwaardering van het grote aantal metacarpalia en falangen betekenen. Daarom werd een gewogen score gemaakt. De bijdrage aan de totale score voor radius en ulna werd daarbij gesteld op 20%, voor metacarpalia en falangen op 30%, en voor carpalia op 50%.



Figuur 5.2. : Bones rated according the Tanner-Whitehouse-Healy-method (1959), in order to estimate the skeletal age.

De som van alle gewogen scores levert een zogenaamde rijpingsscore op tussen 0 en 1.000, welke voor jongens en meisjes in aparte tabellen kan worden vergeleken met een standaard. Op deze wijze kan men uit de rijpingsscore een equivalente skeletleeftijd vaststellen.

Meetprocedure

In ons onderzoek hebben wij zowel op de voortest als op de natest een röntgenfoto van de linkerhand van alle proefpersonen genomen met behulp van de Utilis 15/18, fabrikaat Enraf te Delft. De buisspanning was 45 KV met een expositietijd van 0,3 seconde op een afstand van 80 cm. De palmzijde van de hand werd op een enveloppe film gelegd (Agfa-Gevaert, Osray), waarbij de vingers licht gespreid waren, en de as van de hand, de pols en onderarm zich in een rechte lijn bevonden. Dit levert een plaatselijke belasting op van 30 mR, hetgeen volgens Ziedses des Plantes (1971) zeker drie ordes beneden de gevarengrens blijft, zodat door deze expositie geen biologische invloed op het groeiende bot verwacht mag worden.

De stralingsbescherming van proefpersoon en proefleider geschiedde volgens voorschrift met een loden schort.

De beoordeling van de skeletleeftijd aan de hand van de gemaakte röntgenfoto's geschiedde door een van te voren getrainde beoordelaarster. Alle foto's van zowel voor- als natest werden tegelijk en door elkaar heen gescoord. Voor de betrouwbaarheid van de scoring, zie 6.1.4. reproduceerbaarheid skeletleeftijd.

Verschillen in trainingseffect

Het is voor ons een belangrijke vraag of lichamelijke training dezelfde veranderingen teweegbrengt, en in dezelfde mate bij vroeg rijpende zowel als bij laatrijpende jongens.

Tanner (1962) heeft een analyse gemaakt van de factoren welke de groei en ontwikkeling zouden reguleren. Daarbij benadrukt hij als endogene factor de aanleg. Verschillende exogene factoren zoals ziekte, ondervoeding en lichamelijke inactiviteit kunnen de normale ontwikkeling remmen. Er bestaan echter twijfels of er exogene factoren zijn die de ontwikkeling zouden kunnen stimuleren, met name of lichamelijke training bij kinderen een groter effect zou hebben dan bij volwassenen. Een grotere trainbaarheid bij kinderen zou men mogen verwachten door de hogere concentratie van het groeihormoon in het bloed en van de androgenen tijdens de pubertaire groeispurt. Ekblom (1969) toonde aan dat een duurtraining over een periode van 36 maanden bij jongens in de puberteit een versnelde lengtegroei veroorzaakte.

Een vergelijking van trainingseffecten tussen volwassenen en kinderen is overigens zeer moeilijk omdat er veelal een verschil bestaat in trainingstoestand aan het begin van de training, en omdat het erg moeilijk is om de trainingsarbeid voor volwassenen en kinderen relatief gelijk te houden.

Hettinger (1968) vond dat voor kracht de maximale trainbaarheid voor mannen tussen de twintig en dertig jaar ligt. Volgens Ikai (1969) is het locale spieruithoudingsvermogen het best trainbaar bij jongens tussen twaalf en vijftien jaar.

Eriksson (1972) vond als gevolg van een looptraining dat na zestien weken het aerobe uithoudingsvermogen, gecorrigeerd voor de groei gedurende diezelfde periode, bij jongens tussen elf en dertien jaar een verbetering vertoonde van 14%. Dit komt overeen met trainingsresultaten van volwassenen.

Klissouras e.a. (1973) vonden bij ééneiïge tweelingen na een tien weken durende duurtraining van een van de tweelingen per paar dat de maximale zuurstofopname bij dertienjarigen niet verschilde, maar wel bij de zestienjarige tweelingen. De onderzoekers veronderstelden dat dit verschil in trainbaarheid werd veroorzaakt door het feit dat de dertienjarige tweelingen zich in de groeispurt bevonden. Het is mogelijk dat in die periode de hormonale activiteit optimaal is voor groei en ontwikkeling, zodat extra trainingsstimuli geen waarneembaar (trainings) effect meer hebben.

Daniëls e.a. (1971) konden bij tien- tot vijftienjarige jongens geen verbetering van de maximale zuurstofopname per kilogram lichaamsgewicht aantonen, omdat maximale zuurstofopname en lichaamsgewicht over een periode van tweeëntwintig maanden beide met ongeveer 23% waren toegenomen.

Het lijkt erop dat tijdens de groei de maximale zuurstofopname meer verband houdt met groei en ontwikkeling dan met training.

Bouchard e.a. (1968) vonden dat de biologische leeftijd een grotere invloed had op de lichaamsbouw en lichamelijke prestaties dan de mate van lichamelijke activiteit.

Ook Cerny (1969) veronderstelde dat de ontwikkeling niet sterk beïnvloed wordt door de mate van lichamelijke activiteit blijkens een vijf jaar durend longitudinaal onderzoek bij groepen kinderen

met verschillen in lichamelijke activiteit gedurende die periode.

Resultaten

De kalenderleeftijd van onze proefgroep was binnen zeer nauwe grenzen gelijk. Bij een gemiddelde leeftijd van 12,5 jaar op de voortest is de standaardafwijking 0,4 jaar en het leeftijdsverschil tussen de oudste en de jongste bedraagt slechts 1,8 jaar (zie tabel 2.2.)

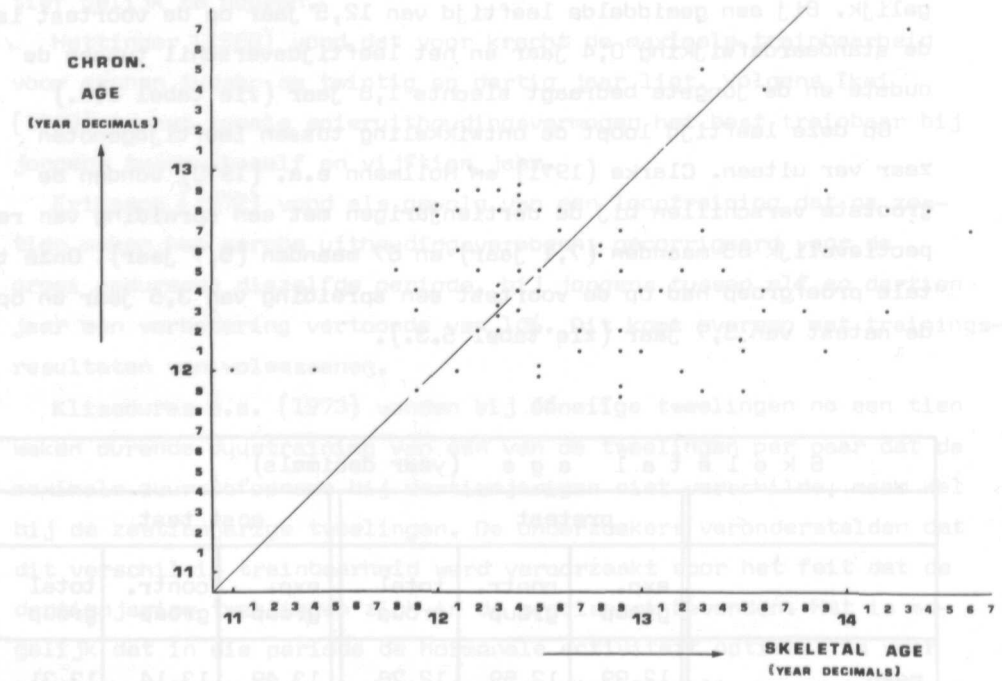
Op deze leeftijd loopt de ontwikkeling tussen leeftijdgenoten zeer ver uiteen. Clarke (1971) en Hollmann e.a. (1970) vonden de grootste verschillen bij de dertienjarigen met een spreiding van respectievelijk 85 maanden (7,1 jaar) en 67 maanden (5,7 jaar). Onze totale proefgroep had op de voortest een spreiding van 3,6 jaar en op de natest van 3,7 jaar (zie tabel 5.3.).

Skeletal age (year decimals)						
	pretest			post test		
	exp. group	contr. group	total group	exp. group	contr. group	total group
mean	12.99	12.59	12.78	13.49	13.14	13.31
stand.dev.	0.82	0.63	0.75	0.80	0.85	0.84
minimum	11.00	11.80	11.00	11.80	11.70	11.70
maximum	14.60	14.40	14.60	14.90	15.40	15.40
range	3.60	2.60	3.60	3.10	3.70	3.70
number	33	37	70	33	37	70

Tabel 5.3. : Skeletal age (year decimals) of experimental, control and total group on pre- and post test.

Uit figuur 5.4. blijkt dat de spreiding in skeletleeftijd het dubbele is van de kalenderleeftijd. Dat evenwel de verschillen niet zo groot zijn als bij Clarke (1971) en Hollmann e.a. (1970) zou verklaard kunnen worden uit het feit dat onze groep op de eerste plaats een op

grond van studieprestaties geselecteerde groep was en zeker geen doorsnede van alle jongens van die leeftijd en op de tweede plaats een kleinere variatie in sociaaleconomisch milieu vertoonde.



Figuur 5.4. : Relation between skeletal age (abscissa) and chronological age (ordinate) of all subjects on pretest (n = 70).

Uit onze resultaten (zie tabel 5.5.) blijkt niet alleen dat de correlaties van skeletleeftijd met antropometrische maten relatief zeer hoog zijn, maar dat er met de kalenderleeftijd geen enkel verband bestaat. Ten aanzien van de fysiologische variabelen vertoonde de W_{170} geen significante correlatie met skeletleeftijd (- 0,12) en de handgrip was de enige eenvoudige prestatietest die een significante correlatie vertoonde met skeletleeftijd (+ 0,52).

De correlaties van zowel antropometrische als fysiologische variabelen met de skeletleeftijd komen overeen met de gegevens van Beunen e.a. (1972) en Borms (1972), die eveneens met 12 en 13-jarige jongens

onderzoek verrichtten.

physiological variables	chronological age	skeletal age	anthropometric variables	chronological age	skeletal age
plate tapping	-.13	-.11	weight	-.02	.70 *
sit and reach	-.07	.04	height	.12	.65 *
50 m. shuttle run	.03	-.10	biacromial diameter	.15	.54 *
bent arm hang	.00	-.04	biiliocristal diameter	.01	.50*
vertical jump	.05	-.03	upperarm circumference	-.09	.45*
handgrip	.00	.52*	wrist breadth	-.05	.52*
12 min. run walk	.01	-.02	corr. upperarm diameter	.05	.54*
W_{170}	.19	-.12	% fat	-.16	.24*

Tabel 5.5. : Spearman correlation coefficients of anthropometric and physiological variables with chronological age and skeletal age of the total group (n = 70) on pretest (* = $p \leq .05$).

De experimentele periode besloeg een periode van 0,8 jaar. Gedurende deze tijd nam de skeletleeftijd toe, variërend van 0,0 jaar tot maximaal 2,1 jaar (zie tabel 5.6.). De toename in skeletleeftijd blijkt geen relatie te vertonen met de skeletleeftijd (zie figuur 5.7.), noch met de kalenderleeftijd op de voortest (zie figuur 5.8.).

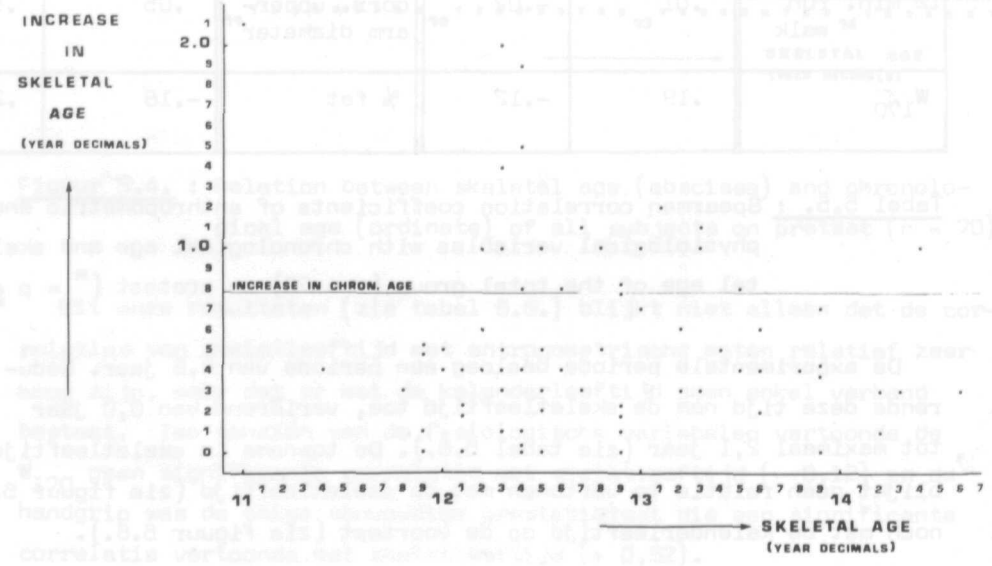
Verwachting

Het lijkt in de verwachting te liggen dat door verschillen in ontwikkeling tussen de jongens uit de proefgroepen het effect van de extra

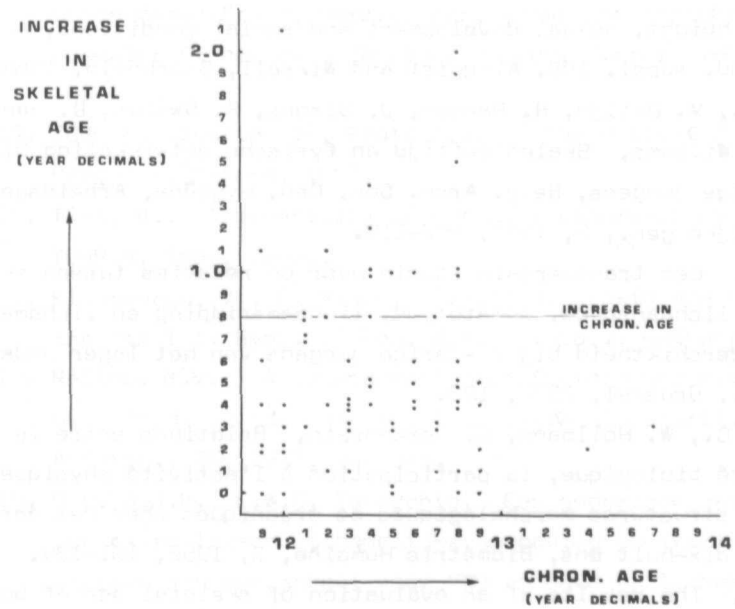
lessen lichamelijke oefening in verschillende mate de afhankelijke variabelen zal beïnvloeden. Daarmee is het opnemen van de skeletleeftijd als indicatie voor de biologische leeftijd zeker gerechtvaardigd.

increase in skeletal age between pre- and post test			
	exp. group	contr. group	total group
mean	0.50	0.55	0.53
stand. dev.	0.40	0.47	0.44
minimum	0.00	0.00	0.00
maximum	2.10	2.00	2.10
range	2.10	2.00	2.10
number	33	37	70

Tabel 5.6. : Increase in skeletal age (year decimals) between pre- and post test of experimental, control and total group.



Figuur 5.7. : Relation between skeletal age on pretest and increase in skeletal age of all subjects (n = 70).



Figuur 5.8. : Relation between chronological age on pretest and increase in skeletal age of all subjects (n = 70).

Literatuur

1. Andersen, E., Skeletal maturation of danish school children in relation to height, sexual development and social conditions, Acta Paed. Scand. suppl. 185, Almquist and Wiksell, Stockholm, 1968.
2. Beunen, G., M. Ostijn, R. Renson, J. Simons, P. Swalus, O. van Gerven, E.J. Willems, Skeletleeftijd en fysische ontwikkeling bij twaalfjarige jongens, Belg. Arch. Soc. Gen. Hygiëne, Arbeidsgen. en gerechtelijke gen., 2, 1972, 102-119.
3. Borms, J., Een transversale studie over de relaties tussen skeletale leeftijd, lichaamsbouw, somatotype, lichaamsrijping en lichamelijke prestatiegeschiktheid bij 12-jarige jongens van het lager onderwijs, Diss. V.U., Brussel, 1972, 120.
4. Bouchard, C., W. Hollmann, G. Herkenrath, Relations entre le niveau de maturité biologique, la participation à l'activité physique et certaines structures morphologiques et organiques chez des garçons de huit à dix-huit ans, Biométrie Humaine, 3, 1968, 101-139.
5. Cerny, L., The results of an evaluation of skeletal age of boys 11-15 years old with different régime of physical activity, Physical Fitness' Assessment, Charles University, Praha, 1969, 56-59.
6. Clarke, H.H., Physical and Motor Tests in the Medford Boy's Growth Study, Prentice-Hall, Englewood Cliffs,1971, 10-47.
7. Cumming, G.R., Correlation of athletic performance and aerobic power in 12 to 17 year old children with bone age, calf muscle, total body potassium, heart volume and two indices of anaerobic power, Proc. of the 4th Int. Symp. of Ped. Work Physiol., Nathanya, 1973, 109-134.
8. Daniëls, J., N. Oldridge, Changes in oxygen consumption of young boys during growth and running training, Med. and Sci. in Sports, 3, 4, 1971, 161-165.
9. Ekblom, B., Effect of physical training on oxygen transport system in man, Acta Phys. Scand., suppl. 328, Stockholm, 1969.
10. Eriksson, B.O., Physical training, oxygen supply and muscle metabolism in 11-13 year old boys, Acta Phys. Scand.; suppl. 384, Stockholm, 1972.
11. Greulich, W.W., I.S. Pyle, Radiographic Atlas of Skeletal Development of the hand and wrist, Stanford, Stanford Univ. Press, 1950.
12. Hettinger, Th., Isometrisches Muskeltraining, Stuttgart, G.Thieme,1968.

13. Hollmann, W., C. Bouchard, Untersuchungen über die Beziehungen zwischen chronologischem und biologischem Alter zu spiroergometrischen Messgrössen, Herzvolumen, anthropometrische Daten und Skelettmuskelkraft bei 8-18 jährigen Jungen, Z.Kreisl.-Forsch., 59, 1970, 160-176.
14. ICRP Publication 16, Protection of the patient in X-ray diagnosis, Pergamon, Oxford, 1969.
15. Ikai, M., Trainability of muscular endurance as related to age, ICHPER, 1969, 28-35.
16. Klissouras, V., G. Weber, Training: growth and heredity, Proc. of the 4th Int. Symp. of Ped. Work Physiol., Nathanya, 1973, 209-215.
17. Malina, R.M., A consideration of factors underlying the selection of methods in the assessment of skeletal maturity, Am. J. Phys. Anthrop., 35, 1971, 341-346.
18. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor VWO en HAVO op lichamelijke en geestelijke groei en ontwikkeling, Werkstuk, Lab. voor Psychofysiologie, Univ. v. Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv., 59, 1971, 543 en 604.
19. Szabo, S., J. Doka, P. Apor, K. Somogyvari, Die Beziehungen zwischen Knochenlebensalter, funktionellen anthropometrischen Daten und der aeroben Kapazität, Schweiz. z. Sportmed., 20, 1972, 109-115.
20. Tanner, J.M., R.H. Whitehouse, M.J.R. Healy, Part I Standards for skeletal age, 1959; Part 2 A new system for estimating skeletal maturity from the hand and wrist, with standards derived from a study of 2600 healthy British children, The scoring system, Paris, Int. Children's Centre, 1962.
21. Tanner, J.M., Growth at adolescence, Blackwell, Oxford, 1962.
22. Wieringen, J.C. van, F. Wafelbakker, H.P. Verbrugge, J.H. de Haas, Groeidiagrammen Nederland 1965, NIPG (TNO), Wolters Noordhoff, Groningen, 1968.
23. Voorstel leerplan Rijksscholen voor lichamelijke oefening voor V.W.O., H.A.V.O. en M.A.V.O., Ministerie van Onderwijs en Wetenschappen, Den Haag, 1968.
24. Wijn, J.W. de, J.H. de Haas, Groeidiagrammen van 1-25-jarigen in Nederland, N.I.P.G., Leiden, 48, 1960.
25. Ziedses des Plantes, B.G., Persoonlijke mededeling dd. 16-8-1960.

5.2. BUITENSCHOOLSE LICHAMELIJKE ACTIVITEIT

De kinderen in ons onderzoek brengen - niet gerekend de tijd dat zij slapen - ongeveer 30% van de tijd op school door. Wanneer wij mogen veronderstellen dat in die tijd de lichamelijke activiteit niet veel zal verschillen, betekent dit dat de wijze waarop de tijd buiten de school wordt doorgebracht van grote invloed kan zijn op de intellectuele en lichamelijke prestaties van het kind. Op de totale hoeveelheid tijd die de leerlingen buiten de school doorbrengen, en waarin zij naar eigen keuze meer of minder lichamelijk actief kunnen zijn, vormen de twee extra uren een zeer gering aandeel. Dit heeft tot gevolg dat, wanneer men te maken zou hebben met kinderen die in hun vrije tijd lichamelijk actief zijn, het effect van deze extra uren - zo het al zou optreden - waarschijnlijk niet meer te achterhalen zal zijn.

Om voornoemde reden is het dan ook noodzakelijk om in het kader van dit onderzoek een indruk te verkrijgen van de mate waarin de leerlingen buiten de school lichamelijk actief zijn, en in welke mate de leerlingen hierin onderling verschillen.

Meetvoorwaarden en toepassingsmogelijkheden

Bij het meetbaar maken van dagelijkse lichamelijke activiteit dient met een aantal feiten rekening te worden gehouden, zoals:

- het te gebruiken instrument moet activiteit kunnen meten zonder dat daarbij het normale levenspatroon van de proefpersoon wordt veranderd (Edholm, 1966);
- een goede meetmethode dient te kunnen worden toegepast bij grote groepen proefpersonen.

Wat betreft het meten zelf dient men voor ogen te houden dat voor het verkrijgen van de juiste informatie ten aanzien van de lichamelijke activiteit in het algemeen geldt dat:

- de periode waarover gemeten wordt niet te beperkt mag zijn, maar tenminste een periode van 24 uur moet beslaan;
- de metingen de gehele week dienen te beslaan, met inbegrip van het weekeinde;
- de meting bij voorkeur herhaald behoort te worden, bij voorbeeld zomer en winter;
- het bij sommige onderzoeken tevens van belang is om informatie te

te verkrijgen over de frequentie, duur en intensiteit van de piek-activiteiten (Edholm, 1966; De Boorder, 1971).

De bestaande meet instrumenten zijn ontwikkeld om met behulp van metingen van de dagelijkse lichamelijke activiteit: a. de arbeidsprestatie van een individu op de juiste wijze te kunnen interpreteren; b. het dagelijkse energieverbruik te berekenen bij studies van het voedselverbruik; c. de relatie vast te stellen tussen lichamelijke activiteit en het voorkomen van hart en vaatziekten (Weiner, 1969; De Boorder, 1971).

Aangezien de methode die wordt toegepast voor het meten van dagelijkse activiteit, sterk wordt bepaald door de vraagstelling van de onderzoeker, dient te worden nagegaan welke van de bestaande meetinstrumenten voor ons onderzoek het meest geschikt zijn.

Methoden voor het meten van lichamelijke activiteit

De methoden die ontwikkeld zijn om de lichamelijke activiteit te kunnen meten, zijn te verdelen in: 1. observatiemethoden; 2. meetmethoden; 3. een combinatie van deze twee (Edholm, 1966).

De meest bekende observatiemethoden zijn: vragenlijsten, dagboeken, 'time and motion' studies en 'cinéphotography'.

Tot de meetmethoden behoren: meting van het zuurstofverbruik, het ademminutenvolume, de hartfrequentie, de voedselopname van de proefpersoon en de pedometer (stappenteller).

De vragenlijsten kunnen op twee manieren worden gebruikt, namelijk: in te vullen door de proefpersoon zelf (Reiff e.a., 1967; Montoye, 1971; Goldsmith e.a., 1971; Buskirk e.a., 1971), of in te vullen door een interviewer na antwoord op vragen (Reiff e.a., 1967; Montoye, 1971, Buskirk e.a., 1971; De Boorder, 1971). De periode waarover gemeten wordt, kan maanden of zelfs jaren beslaan. Over het algemeen heeft men geen betrouwbaarheidsmetingen uitgevoerd (Edholm, 1966).

De wijze waarop vragenlijsten worden gescoord, varieert van het indelen van mensen in categorieën tot gedetailleerde scores op basis van het energieverbruik van de verschillende activiteiten (Durnin e.a., 1967; Durnin, 1971).

De dagboekmethode is een variant op de vragenlijst, waarbij de proefpersoon een dagboek bijhoudt van zijn dagelijkse activiteiten in een bepaalde periode. Aangezien deze methode de medewerking vereist van de proefpersonen, is deze ongeschikt voor toepassing bij

kinderen.

De 'time and motion' studie is een methode waarbij de proefpersonen direct geobserveerd worden. Deze methode kan alleen worden toegepast wanneer een groep proefpersonen zich binnen het gezichtsveld van de observator bevindt. Deze methode is zeer geschikt voor onderzoek tijdens het verblijf van de kinderen op school, maar niet voor de buitenschoolse activiteiten.

De 'cinephotography' is in feite een 'time and motion' studie vanaf film. Deze methode is dan ook zeer geschikt voor het beoordelen van de activiteit van de leerlingen in de klas. Voor de periode na school gelden echter dezelfde bezwaren als reeds genoemd.

Aangezien de vragenlijstmethode eenvoudig toepasbaar is, hebben wij deze in ons onderzoek opgenomen. In verband met de leeftijd van onze proefpersonen is een interviewer ingeschakeld.

De 'meet'-methoden zijn over het algemeen gericht op het verkrijgen van gegevens over het energieverbruik van de proefpersoon. Onze belangstelling gaat echter niet zozeer uit naar de mate van energieverbruik, maar meer naar de duur en intensiteit van de buitenschoolse lichamelijke activiteiten. Hieruit valt namelijk af te leiden in welke mate deze activiteiten een trainingseffect op het lichaam kunnen veroorzaken, welk effect dan niet zou mogen worden toegeschreven aan het deelnemen aan de lessen lichamelijke oefening. Aangezien de voornoemde meetmethoden tevens tamelijk gecompliceerd zijn, komen ze dus niet in aanmerking voor toepassing in het onderhavige onderzoek.

Uitzondering hierop vormt de meting van de hartfrequentie, aangezien deze maat inzicht geeft over de intensiteit van de activiteit voor de betreffende proefpersoon, en het trainingseffect ervan op het cardiorespiratoire systeem.

Het meten van de hartfrequentie met behulp van telemetrische technieken wordt veel toegepast, maar is voor het meten van dagelijkse lichamelijke activiteit ongeschikt in verband met de hoge kosten van de apparatuur, het beperkte afstandbereik van de zender, en het grote aantal proefpersonen.

Daarentegen vormen de hartfrequentietellers (Rowley e.a., 1959; Friedman e.a., 1963; Wolff, 1969; Bradfield, 1971) een categorie meetinstrumenten die op grote schaal en over tamelijk lange perioden zijn toe te passen. De tellers worden geactiveerd door de

R-top van het ECG. Afmeting en gewicht van het instrument zijn zo gering dat het mogelijk is het te dragen zonder daarbij enige hinder te ondervinden.

Dit laatste argument geldt eveneens voor de 'pedometer' of stappenteller, een apparaatje ter grootte van een zakhorloge, dat bewegingen in verticale richting registreert (Zie foto 5.9.).

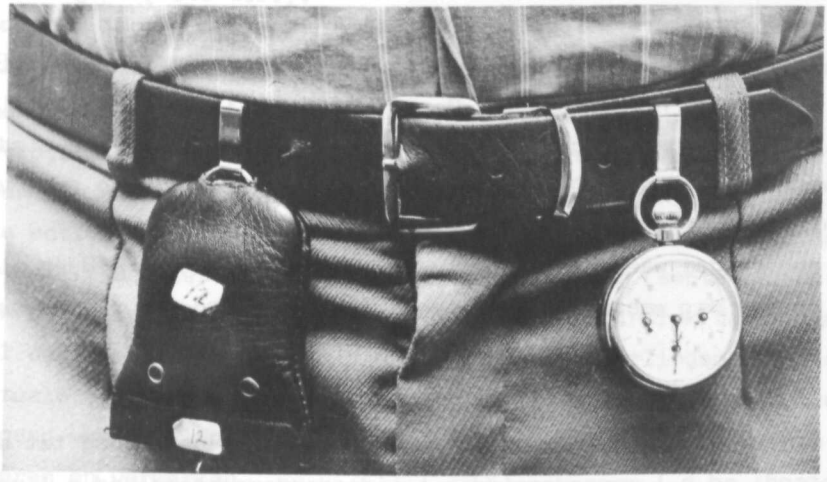


Foto 5.9. : Pedometer, attached to the waist of a subject (right) and covered by a case (left).

Indien bevestigd aan de romp, zal het de verticale verplaatsing van het lichaamszwaartepunt, die bij iedere pas plaatsvindt, afzonderlijk meten. Door de pedometer aan het been te bevestigen en tevens de gemiddelde paslengte van een persoon te bepalen, is het mogelijk om een indruk te krijgen van de afstand die iemand lopend heeft afgelegd (Dorris e.a., 1957; Chirico e.a., 1960; Stunkard, 1960; Stunkard e.a., 1962).

Terwijl volwassenen veelal een zittend, staand, dan wel lopend bestaan leiden, kunnen kinderen zeer actief zijn zonder zich daarbij duidelijk te verplaatsen: bukken, springen, e.d.. Het bevestigen van de stappenteller aan de romp, zonder verdere interpretatie van het aantal geregistreerde verticale bewegingen dan dat het een maat vormt voor lichamelijke activiteit, verdient in dat geval naar onze mening de voorkeur.

Alhoewel de hartfrequentieteller naar voren komt als de beste meetmethode voor het bepalen van dagelijkse lichamelijke activiteit

in ons onderzoek - voornamelijk omdat deze meetmethode behalve inzicht te geven over de totale hoeveelheid, ook iets weergeeft van de intensiteit van de activiteit - is in verband met de kosten van de hartfrequentieteller en de benodigde randapparatuur de keuze toch op de stappenteller gevallen.

Meetprocedure stappentellers

De stappentellers werden in het najaar (november) en in het voorjaar (maart) gebruikt voor het bepalen van de hoeveelheid activiteit van de leerlingen na schooltijd. Om een zo duidelijk mogelijk beeld te krijgen, diende het gehele weekpatroon, inclusief het weekeinde, bestudeerd te worden. Daartoe werden de stappentellers op drie verschillende tijdstippen meegegeven, te weten:

a. op een dag waarop de school pas in de loop van de middag eindigde;
b. op een dag waarop de leerlingen de gehele middag vrij van school waren;
c. tijdens een weekeinde.

De metingen werden uitgevoerd van ad a.) maandagnamiddag tot dinsdagochtend; ad b.) woensdagmiddag tot donderdagochtend; ad c.) vrijdagnamiddag tot maandagochtend.

Om eventuele invloed van dag-tot-dagvariatie op het klassengemiddelde, onder invloed van schoolwerk en weersomstandigheden, zoveel mogelijk te voorkomen, werd voor iedere meting een steekproef uit alle vier klassen genomen. Tevens werden de drie metingen per leerling per meetperiode gespreid over een aantal weken.

De score werd uiteindelijk bepaald door de gemiddelde score te berekenen over de overeenkomstige meettijdstippen in de twee meetperioden.

De totaalscore kwam tot stand door de aldus verkregen scores voor de maandag, woensdag en het weekeinde te sommeren (zie tabel 5.10.).

Gelijktijdig met de eerste meting werd aan de leerlingen een brief meegegeven voor de ouders, waarin vermeld stond dat het hier ging om een meting van de activiteit van de brugklasleerlingen. Tevens werd de ouders verzocht behulpzaam te zijn bij een soepele afwerking van de metingen, door hun kind eraan te helpen herinneren de universeelklok opnieuw te bevestigen wanneer zij van broek zouden verwisselen.

De kinderen werd meegedeeld dat het hier ging om een universeelklok die een aantal lichamelijke aspecten zou meten zoals activiteit, temperatuur, transpiratie en vele andere.

Aangezien het karakter van de meting en de werking van het instrument zoveel mogelijk geheim diende te blijven, werden de stappentellers in een leren etui verpakt waarvan bovendien de opening met een loodje was verzegeld (foto 5.9.).

Meetprocedure vragenlijst

In aanvulling op de stappentellerscores hoopten wij door middel van een vragenlijst een indruk te krijgen in welke mate de verschillende vormen van activiteit bijdragen aan de totale buitenschoolse lichamelijke activiteit. Daartoe werd onze vragenlijst tweemaal afgenomen door dezelfde interviewer, namelijk in het najaar (september/oktober) en in het voorjaar (april/mei), om zodoende een indruk te krijgen van zowel de zomer- als winteractiviteiten van de leerlingen, alsmede om eventuele veranderingen in het activiteitspatroon, als gevolg van de overgang van het lager naar het voortgezet onderwijs, te kunnen constateren en verwerken.

De vragenlijst omvatte vragen over:

- het vervoer van en naar school;
- activiteiten in verenigingsverband;
- activiteiten buiten de verenigingen;
- rookgewoonten (zie bijlage V).

De activiteiten werden uitgedrukt in tijd, het aantal minuten per week dat de leerlingen lichamelijk actief waren. Daarbij werd ervan uitgegaan dat activiteiten zoals liggen, zitten en staan buiten beschouwing konden worden gelaten, gezien het beperkte energieverbruik bij dergelijke activiteiten, respectievelijk 0,0019, 0,0058 en 0,0069 kcal./kg/min (Bink e.a., 1966).

Een sterke stijging van het energieverbruik vindt plaats zodra er actief gewandeld wordt (0,0361 kcal./kg/min). Vandaar dat in dit onderzoek alleen activiteiten met een gelijkwaardig of hoger energieverbruik werden opgenomen in de beoordeling van de buitenschoolse lichamelijke activiteiten van de leerlingen.

Tevens werd rekening gehouden met de neiging tot het overschatten

van de eigen activiteit (Edholm, 1966), door veel navraag te doen over het verzuim van diverse vaststaande activiteiten.

Het gemiddelde aantal minuten activiteit per week, berekend over de twee metingen, vormde de uiteindelijke score.

Verwachting

De verwachting met betrekking tot de invloed van de buitenschoolse lichamelijke activiteit luidt dat, naarmate het activiteitsniveau hoger is, de trainingstoestand hoger, en daardoor de trainingseffectiviteit kleiner zal zijn.

Resultaten

Zoals reeds eerder werd verondersteld, geven zowel de gegevens van de stappentellers als die van de vragenlijst aan dat er grote inter-individuele verschillen zijn in de dagelijkse lichamelijke activiteit van de leerlingen (tabel 5.10. en tabel 5.11.).

Results pedometers during spare time (scores)					
period of measuring	mean	s.d.	min.	max.	range
from Monday afternoon till Tuesday morning	8.000	4.400	2.300	24.300	22.000
from Wednesday noon till Thursday morning	12.000	4.600	2.900	26.600	23.700
from Friday afternoon till Monday morning	31.000	12.600	10.500	65.400	54.900
total score during the three periods	50.800	17.000	17.500	101.600	84.100

Tabel 5.10. : Scores on pedometers, measured during spare time, of 12 and 13-year old boys (n = 70). Each measurement for each boy was done twice during the schoolyear: during autumn (Oct. 1971) and during spring (April 1972).

Om te beginnen is de gemiddelde score op de stappentellers hoger naarmate de tijd waarover gemeten wordt, langer is; de periode van maandag op dinsdag is het kortst (gemiddelde stappenscore: 8.000), en de periode van woensdag op donderdag, met daarin de vrije middag, is weer korter dan het lange weekeinde (respectievelijk 12.000 en 31.000).

De minimaal en maximaal gemeten stappenscore geeft aan dat de verschillen in de individuele activiteit in sommige gevallen het zes-, zeven- of tienvoudige kunnen bedragen.

Results questionnaire habitual physical activity (minutes)						
	mean	s.d.	min.	max.	range	% of total mean
transportation to and from school	151.5	82.4	20.0	400.0	380.0	30.8
organized activities (sports club)	81.5	68.2	0.0	285.0	285.0	16.1
all other activities	279.0	148.5	45.0	785.0	740.0	53.1
total activity during spare time	510.7	180.9	220.0	875.0	655.0	100.0

Tabel 5.11. : Results of a questionnaire concerning physical activity during spare time of 12 and 13-year old boys (n = 70), and measured in the number of minutes spent on physical activity during the week. The questionnaire for each boy was done twice a year, by the same interviewer: during autumn (Sept./Oct.1971) and during spring (April/May '72).

De gemiddelde score op de vragenlijst laat verschillen tussen leerlingen zien tot maximaal een factor vijf. Uit de afzonderlijke categorieën blijkt dat het vervoer van en naar school - in de meeste gevallen per fiets - 30.8% van de totale hoeveelheid tijd die aan lichamelijke activiteit wordt besteed, in beslag neemt.

De activiteiten in verenigingsverband blijken bij deze leeftijdcategorie duidelijk ondergeschikt aan de overige lichamelijke activiteiten van de leerlingen: respectievelijk 16.1% en 53.1%.

De correlatie van de totale score op de stappentellers met de totale score op de vragenlijst was significant, maar laag (0,50). Een verklaring hiervoor zou gevonden kunnen worden in het feit dat fietsen zulk een belangrijke plaats in het activiteitspatroon inneemt, en de stappenteller deze vorm van lichamelijke inspanning niet kan registreren.

Literatuur

1. Bink, B., F.H. Bonjer, H. van der Sluys, Assessment of the energy expenditure by indirect time and motion study, Physical activity in health and disease, ed. K. Evang en K. Lange Andersen, Universitetsforlaget, Oslo, 1966, 207-214.
2. Boorder, Tj. de, Energiebesteding en lichamelijk prestatievermogen, Proefschrift, Universiteit van Amsterdam, Elkerbout, Scheveningen, 1971.
3. Bradfield, R.B., A technique for determination of usual daily energy expenditure in the field, Am. J. Clin. Nutr., 24, sept. 1971, 1148-1154.
4. Buskirk, E.R., D. Harris, J. Mendez, J. Skinner, Comparison of two assessments of physical activity and a survey method for caloric intake, Am. J. Clin. Nutr., 24, sept. 1971, 1119-1125.
5. Chirico, A.M., A.J. Stunkard, Physical activity and human obesity, New England, J. Med., 263, 1960, 935-940.
6. Dorris, R.J., A.J. Stunkard, Physical activity: Performance and attitudes of a group of obese women, Am. J. Med. Sci., 233, 1957, 623-628.
7. Durnin, J.V.G.A., R. Passmore, Energy, work and leisure, Heinemann, Londen, 1967.
8. Durnin, J.V.G.A., Physical activity by adolescents, Acta Paediatr. Scand., Suppl. 217, 1971, 133-135.
9. Edholm, O.G., The assessment of habitual activity, Physical activity in health and disease, Proc. Beitostölen Symp. 1966, ed. K. Evang en K. Lange Andersen, Universitetsforlaget, Oslo, 1966, 187-197.
10. Friedman, M., R.H. Rosenman, A.E. Brown, The continuous heart rate in men exhibiting an overt behavior pattern associated with increased incidence of clinical coronary artery disease, Circulation, 28, 1963, 861-866.
11. Goldsmith, R., T. Hale, Relationship between habitual physical activity and physical fitness, Am. J. Clin. Nutr., 24, dec. 1971, 1489-1493.
12. Montoye, H.J., Estimation of habitual physical activity by questionnaire and interview, Am. J. Clin. Nutr., 24, sept.1971, 1113-1118.
13. Reiff, G.G., H.J. Montoye, R.D. Remington, J.A. Napier, H.L. Metzner

& F.H. Epstein, Assessment of physical activity by questionnaire and interview, Physical activity and the heart, ed.: M.J. Karvonen & A.S. Barry, Springfield, III., Thomas, 1967,chapt. 31.

14. Rowley, D.A., S. Glagov, P.Stoner, Measurement of human heart rate during usual activity, Science, 130, 1959, 976-977.
15. Stunkard, A.J., A method of studying physical activity in man, Am. J. Clin. Nutr., 8, 1960, 595-601.
16. Stunkard, A.J., J. Pestka, The physical activity of obese girls, Am. J. Diseases Childr., 103, 1962, 116-121.
17. Weiner, J.S., J.A. Lourie (eds), I.B.P. Handbook no. 9, Human Biology, a guide to field methods, Blackwell, Oxford, 1969.
18. Wolff, H.S., Socially acceptable monitoring, Ergonomics, 12, 1969, 477.

5.3. LESGEEFGEDRAG VAN DE LERAREN LICHAMELIJKE OEFENING

In een onderzoek naar de relatie tussen lichamelijke oefening en persoonlijkheidskenmerken is het van groot belang de variabele 'leraar' te leren kennen om zodoende iets te kunnen zeggen over de invloed die deze storende variabele op de resultaten van het onderzoek zou kunnen uitoefenen. Vanuit experimenteel oogpunt is het ideaal wanneer leraren onderling niet verschillen: het ideaal van uniforme toediening van de experimentele variabele. Wanneer echter, zoals in dit onderzoek, de lessen lichamelijke oefening, dat wil zeggen, de onafhankelijke variabele, door vier verschillende leraren worden gegeven, is het de vraag of de werkelijkheid dit ideaal wel redelijk weet te benaderen. Immers juist de leraar heeft, als weinig anderen, de gelegenheid zijn persoonlijkheid in het spel te brengen en op die manier de aantrekkingskracht en invloed van de onafhankelijke variabele, meer lessen lichamelijke oefening per week, ten aanzien van de leerlingen direct te beinvloeden.

Wat betreft de inhoud van de lessen is een leerplan opgesteld waarin door middel van een kaartsysteem alle afzonderlijke lessen qua inhoud der lesstof en indeling precies zijn omschreven (zie 3.). De vier leraren lichamelijke oefening kunnen op die manier aan de vier brugklassen gelijke lessen geven volgens een van te voren vastgesteld programma aan de hand van deze kaarten, zodat in ieder geval uniformiteit wat betreft indeling en inhoud van de lessen bestaat.

Het probleem van een verschillende presentatie van dezelfde lesstof blijft echter bestaan. Indien men nu systematische verschillen in de aard van het gebeuren in de klas constateert, kan men trachten na te gaan in hoeverre deze als oorzaak van eventuele verschillen in effect kunnen worden beschouwd.

Cognitieve en emotionele aspecten

Door middel van observaties van lessen die van te voren op de videoband waren opgenomen, is getracht deze eventuele verschillen op te sporen en zowel kwalitatief als kwantitatief meetbaar te maken.

Uit de literatuur (Simon e.a., in: Bergeijk, 1971) blijkt dat observatiesystemen van de onderwijs-leerprocedure zijn te onderscheiden volgens twee aspecten die bij het samenstellen van categorieën een

centrale plaats innemen:

a. systemen, waarin de nadruk wordt gelegd op het sociaal-emotionele klimaat binnen de schoolklas; hierbij wordt bij voorbeeld nagegaan op welke manier de leraar reageert op gevoelsreacties en hoe hij de leerlingen stimuleert (emotioneel aspect);
b. systemen, waarin de nadruk wordt gelegd op het lesgebeuren wat betreft de inhoud, en de procedure die gevolgd wordt om deze inhoud over te dragen en te oefenen: dus de vaktechnische kwaliteiten van de leraar (cognitief aspect), (Van Bergeijk, 1971).

Van beide systemen is in dit onderzoek gebruikgemaakt omdat zowel verschillen in sociaal-emotioneel klimaat, alsook verschillen in lesgebeuren wat betreft inhoudsoverdracht, van belang kunnen zijn bij de interpretatie van de eventuele effecten van extra lessen lichamelijke oefening.

Het vastleggen van de lessen op videoband

Gedurende de experimentele periode gaven de leraren A en D van de controlegroepen 105 lesuren. De leraren B en C van de experimentele groepen gaven ieder 175 lesuren (zie tabel 3.1.). Uit dit totale lessenbestand werden van iedere leraar 15 lessen opgenomen die inhoudelijk verschilden, maar bij de leraren onderling gelijk waren. De opname geschiedde met behulp van een vaste camera in de toestelruimte van het lokaal, vanwaaruit via een groothoeklens de gehele zaal bestreken kon worden. Het spreken van de leraar werd telemetrisch geregistreerd via een loopmicrofoon welke onder zijn kleding met een koordje om de hals op de borst hing. Het geluid van de leerlingen werd daarmee slechts diffuus opgevangen.

Aan de leraren was van te voren meegedeeld dat van het totale aantal opgenomen lessen slechts een beperkt aantal zou worden gebruikt voor het onderzoek, zonder dat een specificatie werd gegeven om welke lessen het zou gaan. Van de vijftien opgenomen lessen werden er drie op de tape gezet voordat het onderzoek naar het lesgeefgedrag begon; dit geschiedde enerzijds om de leraar en zijn klas vertrouwd te maken met de situatie en anderzijds om de filmers te wennen aan de apparatuur. Door storingen van technische aard waren uiteindelijk slechts

tien maal vier lesopnamen bruikbaar voor beoordeling.

5.3.1. VAKTECHNISCHE BEOORDELING VAN HET LESGEEFGEDRAG

Ten aanzien van het cognitief aspect van het lesgeefgedrag vinden wij bij Van Bergeijk (1971) het volgende:

'Bij het cognitief aspect wordt gedacht aan het lesgeefgebeuren voor-
'zover dit betrekking heeft op de didactische inhoud en procedure,
'die gevolgd wordt om deze inhoud over te dragen, te presenteren, te
'oefenen, en in discussie te stellen en wat dies meer zij. Er vindt
'tussen de leraar en de schoolklas een uitwisseling op cognitief ni-
'veau plaats, die het verloop van de les als didactisch gebeuren con-
'stitueert.'

Het doel van een dergelijk beoordelingssysteem in dit onderzoek is tweeledig:

1. in het experiment vormen de lessen lichamelijke oefening de onafhankelijke variabele.
   Op grond van de doelstellingen van de lichamelijke opvoeding betekent dit dat de lessen, indien ze tot resultaten willen leiden, in ieder geval dienovereenkomstig aangeboden moeten worden, en dat ze derhalve alleen van waarde kunnen zijn als men in de vakwereld een positief oordeel kan uitspreken over de vaktechnische aspecten van het lesgeven;
2. in het experiment worden de lessen lichamelijke oefening aan de vier brugklassen door vier verschillende leraren gegeven. Gezien vanuit de ideale experimentele opzet lijkt het daarom wenselijk deze leraren te vergelijken, om zo de eventuele verschillen te kunnen opsporen en aan te wijzen en zodoende de grootte van het leraareffect wat betreft het cognitief aspect te kunnen meten.

Keuze van de beoordelaars

De keuze van de beoordelaars is gevallen op drie inspecteurs en één inspectrice voor lichamelijke oefening, alsmede op een methodiek docent aan een academie voor lichamelijke opvoeding. Zij kwamen uit verschillende delen van Nederland en hun leeftijd liep ver uiteen. Zij werden gekozen op aanwijzing van een instantie die verder geen bemoeienis had

met het experiment (zie voorwoord).

Genoemde functionarissen mogen in de vakwereld van de lichamelijke oefening gezien worden als het kader dat kan beoordelen of een les lichamelijke oefening aan de impliciete normen voldoet en of deze les als voldoende of onvoldoende kan worden beoordeeld.

Keuze van het systeem

Bij de keuze van het beoordelingssysteem werd uitgegaan van het systeem dat gebruikt wordt bij het afnemen van examens in het lesgeven aan schoolklassen. Daarbij moet door de beoordelaars één cijfer worden gegeven. Dit cijfer bestaat in wezen uit de som van een groot aantal (gewogen) eisen, die ten grondslag liggen aan de totale lesgeefbeoordeling. Voor dit onderzoek werd het relevanter gevonden om de afzonderlijke cijfers voor die verschillende eisen te hanteren, zodat een genuanceerder inzicht in de overeenkomsten en verschillen tussen de leraren kon worden verkregen.

Ras (1971) stelde voor om een analyseschema te gebruiken waarbij over korte perioden van vijf minuten telkens een groot aantal verschillende aspecten wordt beoordeeld. De complexe lesgeefsituatie wordt daarmee gestructureerd; men komt tot een meer genuanceerd oordeel over het lesgeven, waarbij men tracht te vermijden:

- dat een laatste indruk doorslaggevend wordt;
- dat een aspect, hetwelk een 'hobby' vormt van de beoordelaar, een overheersende invloed heeft op de totale beoordeling.

Bij het door ons gehanteerde beoordelingsschema zijn wij uitgegaan van dat van Ras (1971), en wel van dat deel dat betrekking heeft op de specifieke eisen van de lessen lichamelijke oefening. Dit schema bestaat uit de volgende hoofdcategorieën:

1. leerstofkeuze;
2. organisatie;
3. houding en optreden.

Aangepast aan dit onderzoek werden de categorieën:

1. leerstof, oefenstof;
2. organisatie;
3. docent.

Voorts hebben wij een aparte categorie toegevoegd:

4. leerlingen.

Het uiteindelijke beoordelingsschema met onderverdelingen in subcategorieën (zie schema 5.12.) is in overleg met de beoordelaars vastgelegd na een proefbeoordeling. In totaal werden tien subcategorieën gebruikt, die vervolgens driemaal werden gescoord in overeenstemming met de door de leraren gebruikte lesindeling:

1. inleiding ± 5 minuten
2. eerste deel ± 15 á 20 minuten
3. tweede deel ± 15 á 20 minuten.

Elke subcategorie werd door een kruisje gewaardeerd op een 5-puntsschaal.

	1	2	3	4	5
1. LEERSTOF/OEFENSTOF (EXERCISES)					
1.1. aangepast aan vaardigheid (adjustment)					
1.2. verantwoorde opbouw (methods)					
1.3. correcties (indiv.,klassik.) (corrections)					
2. ORGANISATIE (ORGANISATION)					
2.1. opstellingen (toestel,leerling) (formation)					
2.2. veiligheidsmaatregelen (safety)					
3. LEERLINGEN (PUPILS)					
3.1. arbeidsintensiteit (working intensity) (m.b.t. uith.verm.)					
3.2. arbeidsvreugde (joy in working)					
4. DOCENT (TEACHER)					
4.1. houding en optreden (attitude)					
4.2. spreken (speech)					
4.3. uitleg en voordoen (explanation,demonstration)					
	1	2	3	4	5

LESONDERDEEL
(PART OF THE LESSON)

1 = slecht (poor)
2 = onvoldoende (insufficient)
3 = voldoende of niet van toepassing (fair or not applicable)
4 = goed (good)
5 = uitstekend (excellent)

Schema 5.12. : Rating scale for lessons of physical education with four main and ten subcategories.

Beoordelingsprocedure

Bij de beoordeling werd iedere keer één serie van vier dezelfde lessen voor alle beoordelaars tegelijk afgedraaid, en wel steeds in een willekeurige volgorde wat betreft de leraren. Vooraf werd de betreffende les wat betreft inhoud en indeling, zoals op de leskaart staat aangegeven, aan de beoordelaars meegedeeld.

Aan het einde van elk lesdeel werd gevraagd om een beoordeling op de tien subcategorieën, zonder dat de mogelijkheid werd gegeven om onderling overleg te plegen. Ook werd tegengegaan dat de beoordelaars door spontane reacties van goed- of afkeuring elkaar zouden beïnvloeden.

Op verzoek van de beoordelaars werd niet de gehele les afgespeeld maar slechts een gedeelte van ieder onderdeel. De observatietijd was afhankelijk van de aard van het lesonderdeel, en in ieder geval zolang dat van iedere beoordelaar de mening vaststond.

De beoordelaars wisten niet welke de experimentele, en welke de controleklassen waren.

Keuze van de lessen

Door een onafhankelijk persoon werden vier lessen gekozen, die wat betreft leerstof verschilden, en ook over de gehele experimentele periode gespreid waren. Deze lessen waren:

1. ringenles (vóór de herfstvakantie)
2. rekstokles (vóór de kerstvakantie)
3. bok- en kopstandles (vóór de paasvakantie)
4. basketballes (vóór de zomervakantie).

Les 3. is een tweede maal beoordeeld voor een betrouwbaarheidsmeting (zie tabel 6.1.).

Resultaten

De scores van de vijf beoordelaars tezamen werden voor iedere leraar afzonderlijk per categorie per les gemiddeld. Vervolgens werden deze waarden nogmaals gemiddeld over de vier lessen. Deze gegevens, tot stand gekomen door scoring op een 5-puntsschaal, zijn in tabel 5.13. vermeld.

category	teacher A			teacher B			teacher C			teacher D			W
	mean	s.d.	rv	mean	s.d.	rv	mean	s.d.	rv	mean	s.d.	rv	
1 EXERCISES													
1.1 adjustment	3.53	0.09	3	3.70	0.41	2	3.18	0.39	4	3.71	0.42	1	.32
1.2 methods	3.13	0.15	3	3.26	0.38	2	2.84	0.35	4	3.45	0.18	1	.34
1.3 corrections	3.00	0.42	3	3.17	0.40	2	2.53	0.48	4	3.47	0.33	1	.55*
2 ORGANISATION													
2.1 formation	3.60	0.36	2	3.56	0.42	3	3.10	0.32	4	3.64	0.32	1	.30
2.2 safety	3.00	0.05	3	3.19	0.35	2	2.80	0.24	4	3.29	0.46	1	.70**
3 PUPILS													
3.1 work intensity	3.65	0.29	1	3.58	0.43	2	3.14	0.52	4	3.43	0.30	3	.46
3.2 joy	3.60	0.19	1.5	3.60	0.28	1.5	3.17	0.33	4	3.58	0.22	2	.15
4 TEACHER													
4.1 attitude	3.51	0.13	3	3.70	0.29	2	2.58	0.57	4	3.77	0.31	1	.65**
4.2 speech	3.08	0.10	3	3.41	0.24	1	2.51	0.31	4	3.28	0.23	2	.45
4.3 explanation a. demonstr.	3.09	0.18	3	3.06	0.48	3	2.51	0.38	4	3.51	9.37	1	.85**
Total	3.32			3.42			2.84			3.51			

Tabel 5.13. : Judgement of professional skill of teaching behaviour of four teachers in physical education: mean, standard deviation (s.d.) and rank order (rv) of the judgements of four lessons of physical education by five judges, given on a 5-poin rating scale (1 = bad; 2 = insufficient; 3 = sufficient or not applicable; 4 = good; 5 = excellent). Kendall's W (W) is a coefficient of concordance, calculated from the rank of the teachers per category and per judge of the four lessons together.

* = $p \leq .05$ ** = $p \leq .01$

Uit deze gegevens blijkt dat:

1. leraar A, B en D zowel op de tien categorieën afzonderlijk, als in totaal over de vier lessen door de beoordelaars als ruim 'voldoende' worden beoordeeld. Leraar C heeft een totaalscore van 2,84 en wordt op vier van de tien categorieën als 'voldoende' beoordeeld;
2. leraar C op alle categorieën de laagste beoordeling krijgt en leraar D op zeven van de tien categorieën de hoogste beoordeling.

Uit de totale score, gemiddeld over de tien categorieën, komt duidelijk naar voren dat de verschillen tussen de leraren A, B en D uiterst gering zijn. Tussen deze leraren bestaat een verschil van ten hoogste 5,7 procent. De beoordeling van leraar C is echter aanzienlijk lager dan van de overige leraren. Dit verschil bedraagt minimaal 16,9 en maximaal 19,0 procent.

Wanneer over de vier lessen samen de leraren gerangordend worden op iedere categorie per beoordelaar, is het mogelijk een concordantie coëfficiënt (Kendalls W) te berekenen (zie 6.1.).

Het blijkt dat van de tien categorieën er drie significant overeenstemmen wat betreft rangorde van de leraren op één-procents niveau, en één op vijf-procents niveau. Dit betekent dat voor die categorieën de rangvolgorde van de leraren niet aan toeval is toe te schrijven, maar dat deze bepaald wordt door andere factoren; in dit geval zou als verklaring genoemd kunnen worden het verschil tussen de leraren wat betreft hun vaktechnische bekwaamheid. Het probleem is echter dat de rangvolgorde alleen niet zoveel zegt over de absolute verschillen tussen de leraren op de verscheidene categorieën.

5.3.2. SOCIAAL-EMOTIONELE BEOORDELING VAN HET LESGEEFGEDRAG

Steunend op onderzoek van Flanders (1970) en van Van Bergeijk (1971) hebben wij gepoogd een instrument te ontwerpen dat het mogelijk maakt om relevante uitspraken te doen over het 'lesgeefgedrag' van de leraren lichamelijke oefening. Het door Flanders ontwikkelde systeem is van toepassing op de schoolklassituatie. Op grond van de overweging dat de les lichamelijke oefening in sterke mate afwijkt van de zogenaamde 'zaakvakken' was het volgens ons noodzakelijk het systeem van Flanders aan deze situatie aan te passen (Stellwag, 1963; De Block, 1965).

De didactische interactie-analyse van Flanders is een middel ter bepaling van het sociaal-emotionele klimaat in de schoolklas. Via interactie-analyse van het verbale gedrag van leraar en leerlingen is het onder meer mogelijk om directief en non-directief lesgeefgedrag te registreren. Directief en non-directief lesgeefgedrag komt in andere onderzoeken vaak overeen met de termen autoritair-democratisch (Lewin e.a., 1939), dominatief-sociaal geïntegreerd (Anderson, 1967), inclusief-preclusief (Cogan, 1958), directief-permissief (Wisp, 1971; Calvin e.a., 1971), en leraargeoriënteerd-leerlinggeoriënteerd (Haigh e.a., 1971; Mc Keachie, 1971).

Raakpunten met ander onderzoek

Uit onderzoek blijkt dat er verschillen zijn te verwachten ten aanzien van leerling variabelen (lichamelijke en/of persoonlijkheidskenmerken) door de invloed van directief of non-directief lesgeefgedrag van de leraar. H.H. Anderson (1967) was een van de eersten die onderzoek op dit gebied heeft gedaan. Hij inspireerde Lewin e.a. (1939) tot verder onderzoek. Zij concludeerden dat leerlingen verschillend reageren op 'democratisch', 'autocratisch' en 'laissez-faire' gedrag van de leraar tijdens de les. Op hetzelfde gebied, maar dan met gebruikmaking van betere observatieinstrumenten, is onder andere onderzoek gedaan door allereerst Withall (1949) die de 'Climate Index' ontwierp, Bales (1959), Medley e.a. (1958) met hun O.S.C.A.R., Amidon e.a. (1967) met de V.I.C.S. en Flanders (1970) met het 'Flanderssysteem'. De observatietechnieken lijken vaak erg veel op elkaar. Het 'Flanderssysteem' heeft echter als extra kwaliteit dat de observatiegegevens door middel

van matrixanalyse geïnterpreteerd kunnen worden. Tausch e.a. (1967) vonden als effecten van strak leidende methoden bij de leerlingen een weinig originele oplossing van problemen, geringe motivatie, frequente frustratiesituaties, lage concentratie, tendens tot verzet en veelvuldige verstoring van de les en neiging tot overname van vooroordelen en conformistisch denken.

Flanders vat eigen onderzoek samen met de constatering dat non-directief lesgeefgedrag de attitude van de leerling ten opzichte van les en leraar verhoogt en prestaties in die lessen verbetert (Flanders, 1970).

Bovengenoemde onderzoekingen werden uitgevoerd bij andere vakken dan de lichamelijke oefening. Wat betreft onderzoek gedaan binnen het kader van de les lichamelijke oefening zijn ook interessante resultaten te vermelden, hoewel het aantal onderzoekingen op dit gebied zeer beperkt is. Finer (1971) vond bij tien en elfjarigen, dat er geen verschil in prestatieverbetering was bij bepaalde gymnastiekoefeningen wanneer directief en non-directief lesgeefgedrag werd vergeleken.

Binnen de context van een hockeyvereniging vond Veen (1969) in een onderzoek bij 12 en 13-jarige jongens dat, naarmate de 'machtsafstand' kleiner was tussen trainer en getrainden, de prestaties in hockey verbeterden; persoonlijkheidsvariabelen veranderden tijdens het onderzoek niet.

Men moet zich realiseren dat de begrippen directief en non-directief niet eenduidig worden gehanteerd. Ze kunnen betrekking hebben op
a. het leraargedrag, zoals in het onderzoek van Becker (1967) - niet in de les lichamelijke oefening - en
b. op de lesinhoud (programma, methodiek en dergelijke), zoals in de onderzoeken van Veen (1969) en Finer (1971). Een leraar zou in een les die wat betreft inhoud directief is, toch non-directief lesgeefgedrag kunnen vertonen.

Mede in verband met het bovenstaande moet er onderscheid gemaakt worden tussen de volgende onderzoekdesigns:

1. De leraar krijgt na een bepaalde periode of direct in het begin van de experimentele fase van een onderzoek de instructie om te veranderen van lesgeefgedrag (Becker, 1967; Madsen e.a., 1968);
2. De leraar wordt geselecteerd op grond van observatiegegevens en/of

persoonlijkheidskenmerken die gerelateerd zijn aan directieve en non-directieve aspecten van zijn gedrag (Peck, 1959; Barr, 1955; Hunt, 1956; Symonds, 1955; Pankratz, 1967).

In dit onderzoek waren de lessen, gegeven door de vier leraren, voor wat betreft de indeling en inhoud gelijk en werd de leraar in afwijking van bovenstaande designs niet op de een of andere wijze gemanipuleerd. In tweede instantie werd alsnog besloten de sociaal-emotionele aspecten van het leraargedrag te gaan onderzoeken, zij het dan in exploratieve zin: de enige mogelijkheid die nog restte. Wanneer mocht blijken dat er inderdaad duidelijke verschillen bestaan tussen de vier leraren, dan zal met behulp van de resultaten uit de bovengenoemde onderzoeken op dit gebied getracht worden resultaten uit eigen onderzoek te relateren aan de verschillen in onderwijsstijl tussen de leraren.

Interactie binnen de schoolklas

Alle observatiesystemen worden onder meer gekenmerkt door een zekere mate van subjectieve interpretatie van de gegevens die men via de observatie verkrijgt en verwerkt. Het systeem van Flanders (1970) wordt gekenmerkt door 'low-inference', dat wil zeggen dat de waarnemingscategorieën betrekking hebben op specifieke, welomschreven, relatief objectief observeerbare gedragsvarianten (bij voorbeeld het aantal malen dat iemand zegt: 'goed zo', of 'uitstekend'). In tegenstelling met 'low-inference' kent men systemen die gekenmerkt worden door 'high-inference', bij voorbeeld de items waaruit 'rating-scales' zijn opgebouwd: 'enthousiasme', 'behulpzaamheid', etc., welke moeten worden afgeleid uit een reeks gebeurtenissen of situaties. Het 'low-inference'-aspect is in ons systeem gehandhaafd, en waar dit noodzakelijk leek, nog versterkt door enige categorieën te splitsen.

Flanders' systeem is een affectief systeem, hetgeen inhoudt dat de gedragscategorieën refereren aan de invloed die van het gedrag van de leraar uitgaat op het emotionele klimaat van de schoolklas. Natuurlijk impliceert deze benadering reductie: het totaal van het gebeuren registreren is onmogelijk.

Het aspect uit het totaalgebeuren waar Flanders zich op richt, is

de rol die de leraar in de schoolklas vervult. Het gaat om sociaal-emotionele aspecten van zijn gedrag, en om zijn sociale vaardigheid, dit wil zeggen de structurering die hij in sociaal en didactisch opzicht binnen het kader van het schoolklasgebeuren aanbrengt, alsmede de beïnvloedingstendenzen die zijn optreden kenmerken.

Uitgaande van de structuur van een schoolklas bestaan volgens Flanders een drietal gezichtspunten die in voldoende mate toereikend zijn ter verklaring en predictie van een groot deel der gedragingen die in de schoolklassituatie kunnen worden waargenomen:

a. Gezag, dat wil zeggen in hoeverre de leraar het gezag monopoliseert en de leerlingen buiten de besluitvorming houdt. Of in hoeverre hij delegeert en de leerlingen bij de besluitvorming betrekt. In de praktijk zal het leraargedrag zich meestal bewegen tussen beide extremen;

b. Doeloriëntatie, dat wil zeggen in hoeverre de leerlingen weten wat van hen wordt verwacht. Zijn ze in voldoende mate op de hoogte van het doel van de lessen? Begrippen die bij deze problematiek een rol spelen zijn: doelstellingsduidelijkheid versus doelstellingsambiguïteit, met als resultante: doelstellingsattractiviteit. Te veel doelstellingsambiguïteit zal bij de leerlingen resulteren in frustratie en voornamelijk negatieve reacties opwekken;

c. Sociale speelruimte, dat wil zeggen in hoeverre bestaan er mogelijkheden tot het aangaan van sociale interactie en het uiten van ideeën en gevoelens.

Om dit beknopte overzicht van een deel van Flanders' theorievorming te completeren, is het noodzakelijk om nog enige begrippen te introduceren:

d. Flexibiliteit. Schoolklassen waarin het gebeuren door een hoge flexibiliteit wordt gekenmerkt, onderscheiden zich van klassen met een lage flexibiliteit door een veelvuldiger wisseling en gedifferentieerder patroon van gezagsstructuren, doeloriëntatiestructuren en sociale speelruimtestructuren;

e. Afhankelijkheid. Onder afhankelijkheid wordt verstaan de neiging van de leerling om het initiatief aan de leraar over te laten.

<u>De tien categorieën uit het Flanderssysteem</u>

FLANDERS' INTERACTION ANALYSIS CATEGORIES		
TEACHER TALK	RESPONSE	1. ACCEPTS FEELING Accepts and clarifies an attitude of a pupil in a non-threatening manner. Feelings may be positive or negative. Predicting and recalling feelings are included; 2. PRAISES OR ENCOURAGES Praises or encourages pupil action or behaviour. Jokes that release tension, but not at the expense of another individual; nodding head, or saying 'Um hum', or 'Go on' are included; 3. ACCEPTS OR USES IDEAS OF PUPILS Clarifying, building or developing ideas suggested by a pupil. Teacher extensions of pupil ideas are included, but as the teacher brings more of his own ideas into play, shift to category five;
		4. ASKS QUESTIONS Asking questions about content or procedure, based on teacher ideas, with the intent that a pupil will answer;
	INITIATION	5. LECTURING Giving facts or opinions about content or procedures; expressing his own ideas, giving his own explanations, or citing an authority other than a pupil; 6. GIVING DIRECTIONS Directions, commands, or orders to which a pupil is expected to comply; 7. CRITICIZING OR JUSTIFYING AUTHORITY Statements intended to change pupil behaviour from nonacceptable to acceptable pattern, bawling someone out, stating why the teacher is doing what he is doing, extreme self-reference;
PUPIL TALK	RESPONSE	8. PUPIL-TALK, RESPONSE Talk by pupils in response to teacher. Teacher initiates the contact or solicits pupil statement or structures the situation. Freedom to express own ideas is limited;
	INITIATION	9. PUPIL-TALK, INITIATION Talk by pupils which they initiate. Expressing own ideas; initiating a new topic; freedom to develop opinions and a line of thought, like asking thoughtfull questions; going beyond the existing structure;
SILENCE		10. SILENCE OR CONFUSION Pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

Schema 5.14. : The ten categories of Flanders' Interaction Analysis system, in which a distinction has been made in teacher talk (response and initiation) and pupil talk (response and initiation).

Het Flanderssysteem is het best te karakteriseren als een interactie-analyse, en in dit geval betreft het de registratie en analyse van de verbale interactie tussen leraar en leerling aan de hand waarvan het gedrag van de leraar en de leerling uiteindelijk als volgt kan worden gekenschetst:

a. directief gedrag. Dit werkt maximaliserend ten aanzien van de deelname der leerlingen, dat wil zeggen dat voor initiatief en activiteit van de leerlingen weinig speelruimte wordt geboden. Dit gedrag komt vooral tot uiting tijdens het docerend lesgeven, bedoeld als presentatie van de leerstof: 'didactisch éénrichtingverkeer' (onderscheiden door de categorieën 5 t/m 7);

b. non-directief gedrag. Dit werkt minimaliserend ten aanzien van de deelname van de leerlingen, dit wil zeggen dat er voor initiatief en activiteit van de leerlingen speelruimte geboden wordt (onderscheiden door de categorieën 1 t/m 4);

c. leerlinggedrag. Initiatief van de leerling, of antwoord op de vragen van de leraar (onderscheiden door de categorieën 8 en 9).

De tien categorieën die door Flanders in zijn systeem worden onderscheiden worden vermeld en verklaard in schema 5.14.. Ze zijn gedeeltelijk afkomstig uit factoranalyse van lesprotocollen. Zie voor deze factoranalyse tevens: Ryans (1952, 1957), Bartlett (1959), Cosgrove (1959), Medley e.a. (1959), Solomon (1963), McKeachie (1961), Bookhout (1967), Barr (1967).

Ontwikkeling van een instrument en een methode van beoordeling

Voordat met het onderzoek naar beoordelingen van de sociale en emotionele aspecten van leraargedrag kon worden begonnen, was het noodzakelijk om te komen tot een aanpassing van het instrument van Flanders (zie schema 5.14.) en een training van beoordelaars in het hanteren van dit instrument, opdat een redelijke overeenstemming tussen hen zou worden bereikt. Onder leiding van twee psychologen verliep het proces als volgt:

1. Om gedrag van anderen juist waar te nemen en te interpreteren, werd voor de training van de beoordelaars gebruikgemaakt van een counseling pretest, overgenomen uit E.H. Porter (1965). Met behulp van deze methode is het mogelijk om bepaalde gedragingen te onderkennen, zoals 'evaluative', 'interpretive', 'supportive', en 'probing of understanding'. Zodoende werden de beoordelaars geoefend in het onderscheiden van verschillende aspecten in een gesproken of geschreven tekst;
2. Vervolgens werden protocollen (uitgeschreven delen van een gymnastiekles) van de vier verschillende leraren bekeken, met als uitgangspunt het categorieënsysteem van Flanders;
3. Op grond van de punten 1. en 2. kwamen voorstellen ter tafel met als doel een voor de les lichamelijke opvoeding adequaat observatieinstrument te ontwikkelen, steeds met de interactieanalyse van Flanders als uitgangspunt. De resultaten zijn vermeld in het gemodificeerde systeem in schema 5.15.;
4. Gedeelten van de lesprotocollen werden in tijdseenheden verdeeld waarna gepoogd werd het gemodificeerde observatiesysteem te laten toepassen door de beoordelaars;
5. Nadat een zekere communis opinio was bereikt over het te hanteren categorieënsysteem, werden verscheidene proefobservaties gehouden. Iedere beoordelaar scoorde delen van video-opnamen van de verschillende lessen lichamelijke oefening. Deze lessen werden later niet opnieuw gebruikt voor onderzoek. Deze fase werd voortgezet tot een redelijke overeenstemming was bereikt. Om technische en organisatorische redenen kon de training tot een volledige onderlinge overeenstemming tussen de beoordelaars niet plaatsvinden. In het verdere verloop van het onderzoek is getracht dit manco zoveel mogelijk te herstellen (zie 6.1.3.).

MODIFIED FLANDERS' INTERACTION ANALYSIS SYSTEM			
TEACHER	RESPONSE	1. Acceptatie vangevoelens, gedachten en daden, aanmoedigingen en het prijzen, wanneer dit alles collectief is gericht;	Acceptance of feelings, collectively;
		2. Idem als 1., maar individueel gericht;	like 1., but individually;
		3. Meespelen van de leraar zonder verbaal gedrag;	joining a game or performance;
		4. Hulp verlenen zonder verbaal gedrag;	giving aid, non verbal;
		5. Het stellen van 'open' vragen;	asking 'broad' questions;
	INITIATION	6. Het stellen van 'gesloten vragen;	asking 'narrow' quest.;
		7. Instructie, docerend lesgeven en correctie terwijl leerlingen luisteren zonder lichamelijk bezig te zijn;	instructions, while pupils are inactive;
		8. Voordoen door leraar zonder verbaal gedrag;	showing a technique, non verbal;
		9. Instructie, docerend lesgeven en correctie terwijl leerlingen een opdracht uitvoeren;	instruction, while pupils are active;
		10. Het geven van algemene gedragsaanwijzingen;	giving directions;
		11. Expliciete stimulatie;	explicit stimulation;
		12. Critiseren en afwijzen van gevoelens, gedachten en daden, wanneer deze collectief gericht zijn;	criticizing collectively;
		13. Idem als 12, maar individueel gericht;	criticizing, individually;
PUPIL		14. Leerling is bezig, voert opdrachten uit zonder verbale tussenkomst van de leraar;	action and performance;
		15. Voordoen door de leerling en antwoord geven op 'gesloten vragen van de leraar;	showing a technique answering 'narrow' questions;
		16. Initiatief in de meest ruime zin en geven van antwoord op 'open' vragen die door de leraar worden gesteld;	initiative and answering 'broad' questions;
		17. Restcategorie voor stilte en verwarring.	silence or confusion.

Schema 5.15. : The 17 categories of the modified Flanders' Interaction Analysis system.

Verklaring van de categorieën en de verschillen met het Flandersysteem

LERAARGEDRAG Non-directieve beïnvloeding

Categorie 1. Acceptatie van gevoelens, gedachten en daden; aanmoedigen en prijzen; collectief gericht wil zeggen: ten opzichte van de groep als geheel. Bij voorbeeld 'Jullie hebben goed je best gedaan', 'Ik begrijp dat iedereen erg moe is na deze oefening, maar...'. Wat prijzen en aanmoedigen betreft: dit kunnen losse woorden zijn, zoals 'Goed', 'Uitstekend', enz..

Categorie 2. Hetzelfde als categorie 1., maar dan individueel gericht. Bij voorbeeld 'Prima gedaan, Klaas', of 'Dat is een goede opmerking, Kees'.

Naar onze mening is het door Flanders gemaakte onderscheid tussen de categorieën 1., 2. en 3. uit zijn systeem niet essentieel voor de vraagstelling in dit onderzoek, waarin het vooral gaat om leraren duidelijk van elkaar te onderscheiden. Dat we wel een onderscheid hebben gemaakt tussen collectief en individueel, vindt zijn oorzaak in het volgens ons belangrijke psychologische verschil. Immers collectiefgerichte prijzende en/of critische opmerkingen kunnen door de individuele leerling als niet op hem persoonlijk gericht worden opgevat. Tevens dient men in overweging te nemen dat gymnastieklessen over het algemeen vrij collectief zijn opgezet (grote zalen, algemene instructies, e.d.). Daarom zijn hier in het bijzonder individuele opmerkingen indicatief voor de mate van aandacht ('gevoel') van de leraar voor de leerling(en).

Categorie 3. Meespelen zonder verbaal gedrag van de leraar. Deze categorie is opgenomen vanwege de relevantie in de les lichamelijke oefening. Bij voorbeeld de leraar doet mee aan een basketbalspel of een oefening.

Categorie 4. Hulpverlenen zonder verbaal gedrag van de leraar. Voor wat betreft de afwijking van het Flanderssysteem, zie onder cat. 3.. Bij voorbeeld de leraar helpt bij het bokspringen, of hij helpt bij het uitzetten van de toestellen.

Dat categorie 3. en 4. in dit systeem behoren tot het non-directieve leraargedrag vindt zijn rechtvaardiging in het feit dat de leraar in deze twee situaties de (sociaal-emotionele) afstand tussen

zich zelf en de leerling(en) tot een minimum reduceert.

Categorie 5. Het stellen van zogenaamde open vragen: in afwijking van Flanders hebben wij het vragenstellen gesplitst in 'open' en 'gesloten' vragen. 'Open' vragen vallen onder non-directief gedrag omdat ruimte wordt gelaten voor leerling-initiatief: meerdere antwoorden op een vraag kunnen juist zijn. Bij voorbeeld 'Wie weet een variatie op deze oefening (spel)?'.

LERAARGEDRAG Directieve beïnvloeding

Categorie 6. Het stellen van 'gesloten' vragen. Deze worden gesteld met de bedoeling een eenduidig antwoord te krijgen en beperken als zodanig de mogelijkheden tot initiatief bij de leerlingen. Zie ook de opmerkingen onder cat. 5., bij voorbeeld 'Hoeveel keer ben jij over de bok gesprongen?'.

Categorie 7. Instructie, docerend lesgeven en correctie terwijl de leerlingen luisteren, zonder lichamelijk actief bezig te zijn. Deze categorie is enigszins afwijkend van de desbetreffende categorie in het Flanderssysteem, vanwege de nadruk op de inactiviteit van de leerlingen, hetgeen wederom zinvol is in de les lichamelijke oefening (zie cat. 9.). Bij voorbeeld 'Dadelijk gaan we touwklimmen en dat doen we als volgt...', of 'Je moet je benen sluiten bij het zwaaien'.

Categorie 8. Voordoen door de leraar zonder verbaal gedrag. Bij verbaal gedrag: zie de desbetreffende categorie. Dit voordoen is duidelijker in de les lichamelijke oefening dan in de 'leervakken'. Bij voorbeeld met het voordoen van een bepaalde oefening of techniek.

Categorie 9. Instructie, docerend lesgeven en correctie terwijl de leerlingen een opdracht uitvoeren. Het onderscheid met categorie 7. legt de nadruk op de mate van lichamelijke activiteit van het merendeel der leerlingen. Bij voorbeeld 'Probeer sneller te lopen'.

Categorie 10. Het geven van algemene gedragsaanwijzingen. Alle verbale uitingen, die niet in direct verband staan met het instructieve deel van de les. Bij voorbeeld 'De banken tegen de wandrekken aan', of 'In deze les is het zo, dat wanneer ik spreek, jullie je mond houden'.

Categorie 11. Expliciete stimulatie ('pep talk') met referentie aan een extern criterium. Bij voorbeeld 'Kom op jongens, even volhouden, we zitten niet in een bejaardentehuis'.

Naar onze mening bestond de noodzaak om opmerkingen van de leraar, die een duidelijk stimulerend karakter hebben, afzonderlijk te scoren, omdat ze moeilijk zijn onder te brengen in een andere categorie.

Categorie 12. Critiseren en afwijzen, collectief gericht, van gevoelens, gedachten en daden. Zie ook de opmerkingen bij de categorieën 1. en 2.. Bij voorbeeld 'Dit is nu al de derde keer dat ik jullie moet waarschuwen', of 'Van jullie had ik wel een beetje meer sportiviteit verwacht'.

Categorie 13. Hetzelfde als categorie 12., maar individueel gerichte critiek en afwijzing van gevoelens, gedachten en daden. Bij voorbeeld 'Jan, nogmaals, ga zitten'.

Ten aanzien van de categorieën 12. en 13. zie ook de opmerkingen onder de categorieën 1. en 2..

LEERLINGGEDRAG

Categorie 14. Leerling is bezig, voert opdrachten uit, zonder tussenkomst van de leraar. Bij verbaal of ander gedrag van de leraar wordt gescoord op de desbetreffende categorie. Het is een aparte categorie vanwege de relevantie van lichamelijke activiteit van de leerlingen in de les lichamelijke oefening. Het gedrag dat categorie 14. beschrijf valt bij Flanders onder de restcategorie stilte of verwarring. Wij achten het equivalent in de les lichamelijke oefening verschillend in die zin, dat de activiteit zonder tussenkomst van de leraar zinvol geïnterpreteerd kan worden.

Categorie 15. Voordoen van een oefening en het antwoord geven op 'gesloten' vragen door de leerling op verzoek van de leraar. Indien de leraar spreekt, zie de betreffende categorie. Dit is een afzonderlijke categorie omdat het leerlinggedrag niet scoorbaar is onder een andere categorie. In beide gevallen respondeert de leerling op initiatief van de leraar.

Categorie 16. Initiatief van de leerlingen in de meest ruime zin, en het geven van antwoord op 'open' vragen 'zie cat. 5. en 6.). Omdat

het hier alleen 'open' vragen betreft, valt het onderscheid weg dat Flanders maakt in de categorie 8. en 9.. Bij voorbeeld 'Meneer, gaan we vandaag voetballen?'. Als antwoord op de 'open' vraag in categorie 5.: 'Laten we met twee ballen gaan trefballen'.

RESTCATEGORIE

Categorie 17. Stilte en verwarring voorzover het niet kan worden gescoord onder een andere categorie. Bij voorbeeld: Het materiaal vertoont gebreken en moet even hersteld worden.

Keuze van de lessen

De acht lessen van de vier leraren die beoordeeld werden, hadden de volgende inhoud:

1. basketball dribble, ringen duikelen, trefbal;
2. rekstok; (tot de herfstvakantie)
3. hindernisturnen met behulp van bruggen;
4. ladders, verplaatsen in hang; (tot de kerstvakantie)
5. kastspringen;
6. kopstand met behulp van matten; (tot de paasvakantie)
7. basketball spelen;
8. toestel straat (tot de zomervakantie).

Keuze van de beoordelaars

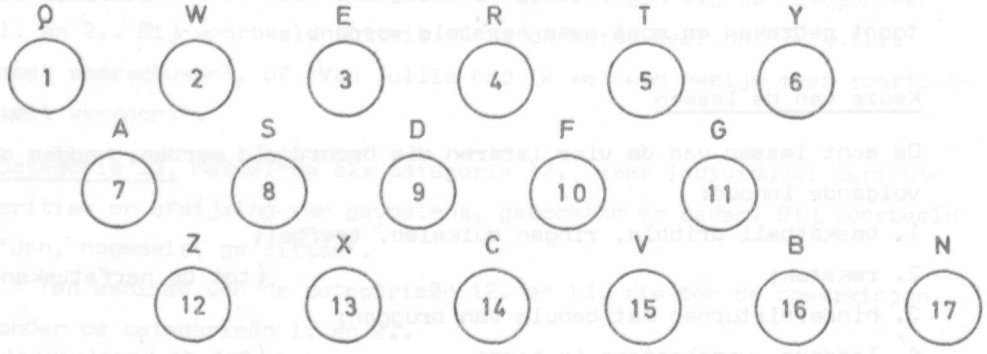
Van de vijf beoordelaars die in leeftijd varieerden van twintig tot vijfentwintig jaar, waren er drie afgestudeerd aan de Academie voor lichamelijke opvoeding te Amsterdam. Ieder van hen had enige jaren onderwijservaring. De andere twee waren respectievelijk derde- en vierdejaarsstudenten aan dezelfde Academie.

Tijdens het onderzoek scoorden de beoordelaars onafhankelijk van elkaar.

Beoordelingsprocedure

In plaats van het 'turven' van bepaalde gedragingen van de leraar en/of leerlingen, ter codering van didactische interactie, zoals dit onder meer in het Flanderssysteem wordt gedaan, is in dit onderzoek gebruikgemaakt van een computer. De 17 categorieën kregen op de

teletype ieder een letterteken van het toetsenbord en werden zodanig gegroepeerd dat de beoordelaar snel en accuraat kon scoren. Met behulp van dopjes waarop de nummers van de categorieën waren aangebracht, werden de toetsen van de teletype afgedekt. Dit is weergegeven in figuur 5.16.



Figuur 5.16. : Keyboard of the teletype with the numbers of the 17 categories indicated on keys.

Bij het observeren van de interactie tussen leraar en leerling in de lessen ging men als volgt te werk: de beoordelaar hoort via de koptelefoon en ziet op een monitor wat door leraar en leerlingen gezegd en gedaan wordt; hij drukt de toets in van de categorie welke hij van toepassing acht op het gebeuren. Slechts wanneer een verandering optreedt in het geobserveerde gedrag van de leraar of de leerlingen, drukt de beoordelaar kortstondig op de toets die het betreffende gedrag representeert. De computer blijft dit gedrag (de bepaalde categorie) dan zolang registreren tot een andere toets wordt ingedrukt. In tegenstelling tot Flanders is het waarnemingsinterval niet 3 sec. maar 1 sec.. De keuze van het 1 sec.-interval berust op de overweging dat nu ook de zeer korte 'statements' van leraar en leerling gescoord kunnen worden, waardoor het accent meer kan vallen op het aantal statements. Op deze manier wordt een zeer nauwkeurige registratie van

overgangen mogelijk naar andere categorieën. Wanneer een les in de boven omschreven vorm in code is omgezet, kan men onmiddellijk daarna beschikken over een door de computer gerangschikte reeks cijfers die terwille van de bruikbaarheid en overzichtelijkheid via een bepaald programma in een matrix zijn ondergebracht. Dit is weergegeven in schema 5.17..

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
1	6									2				3				1
2		56		2					12	2				14	1			2
3																		3
4		1		42			7		18	1				2				4
5		1			86	1	2			2					1	1		5
6						7	1			2				1	0	1		6
7		1			2	1	468	6		18		1	1		22	3		7
8							6	11	1									8
9		8		14				1	228	2	4			60		1		9
10	1				4	1	19			321	2			21	1	5		10
11		2							3	1	15			8		1		11
12							1					7		1				12
13										1			4	1				13
14	3	17		5	2	2	1		51	15	9	1	1	601		7		14
15	1			7			17								33			15
16		1		1	1		1		5	7				3		39		16
17										1							6	17
Tot.	11	87	0	71	24	12	523	18	318	375	30	9	6	715	58	58	7	Tot.
%	0.5	3.6	0.0	3.0	3.9	0.5	21.9	0.8	13.3	15.7	1.3	0.4	0.3	29.9	2.4	2.4	0.3	%

Schema 5.17. : Matrix with steady state cells (diagonal, from upper left to right below) and transition cells: example of an analysis of an arbritary lesson of physical education, with a duration of 40 minutes.

Beschrijving van de matrix

De matrix is samengesteld uit 289 cellen. Van deze cellen zijn er zeventien die op de diagonaal liggen lopend van linksboven naar rechtsonder. Ze worden constantiecellen genoemd. In constantiecellen is het aantal seconden geregistreerd ('steady state cells'); ze representeren de zeventien waarnemingscategorieën van het boven beschreven systeem. De overige cellen, liggend ter weerszijde van de diagonaal, worden overgangscellen genoemd ('transition cells'). Ze registreren de overgangen van de ene categorie naar de andere. Elke cel afzonderlijk kan worden benoemd door eerst gebruik te maken van de rijaanduiding (horizontaal) en daarna de kolomaanduiding (verticaal), identiek aan het verloop in de tijd. In de totaalrij vindt men de absolute totalen van de kolommen, die in de procentenrij bovendien nog in percentages van het totale lesgebeuren kunnen worden weergegeven (schema 5.17.).

Het interpreteren van de matrix

Bij het interpreteren van de matrix kan men aan de hand van de kolompercentages aangeven welk aandeel de betrokken categorie in het totale interactieproces heeft gehad. Een belangrijk gegeven dat uit de matrix kan worden afgeleid, wordt door Flanders de I/D-ratio genoemd.

a. Deze ratio is een quotiënt waarin de mate van directief en nondirectief leraarsgedrag wordt aangegeven. In het systeem wordt de ratio gevormd door de som van de categorieën 1 tot en met 5 te delen door de som van de categorieën 6 tot en met 13. Op deze manier is het steeds mogelijk om de leraren te vergelijken.
   In afwijking van door Flanders gebruikte gegevens uit de matrix is in dit onderzoek gebruikgemaakt van een aantal andere maten, relevant voor de vraagstelling in dit onderzoek, namelijk:
b. Vergeleken worden de frequentietotalen van de categorieën 1 en 2 (acceptatie, prijzen en aanmoedigen - collectief en individueel -) versus de frequentietotalen van de categorieën 12 en 13 (afwijzen, kritiek - individueel en collectief). Dat we hier gebruikmaken van de frequentietotalen en niet van de kolom totalen (zoals bij de I/D-ratio), vindt zijn rechtvaardiging in de overweging dat het bij dit leraargedrag vooral gaat om het aantal malen dat het

plaatsvindt. Hiertoe trekt men het constantiecelaantal per categorie af van het kolomtotaal. Aldus houdt men het frequentietotaal over. De categorieën waarom het hier gaat, betreffen gedrag dat in de praktijk voornamelijk uit korte 'statements' bestaat, zodat de relevantie van de constantiecelinhoud bij dit gedrag minder duidelijk is dan bij de overige categorieën.

c. Per leraar worden de leerlinginitiatieven vergeleken door middel van de kolomtotalen van de categorie 16. De mate van leerlinginitiatief geeft een beeld van de mogelijkheden die de leraar biedt om, wel of niet gewaardeerd, spontaan gedrag te uiten.

d. De totaalmatrix van iedere leraar (gemiddeld over de acht lessen) zal onderzocht worden om na te gaan of er tussen hen opvallende verschillen bestaan. Getracht zal worden om eventueel optredende verschillen zinvol te verklaren.

De hierboven beschreven 'maten' (behalve d) kunnen worden beschouwd als sociaalemotionele graadmeters, omdat ze alle direct iets zeggen over de mate van nondirectief gedrag dat door de leraren wordt vertoond. Dat wij ze belangrijk vinden bij de interpretatie van de matrix vindt zijn rechtvaardiging in de overweging dat:

1. Lessen lichamelijke oefening al zo collectief zijn opgezet, met veel instructie en correctie, in grote zalen, dat juist binnen déze context opmerkingen en/of gedragingen die getuigen van aandacht en 'gevoel' van de leraar voor de leerling, de moeite van het registreren waard zijn en een bijzonder accent verdienen;
2. Andere gegevens die men uit de matrix zou kunnen verkrijgen, zoals de mate van activiteit in de les, zullen in 5.3.3. op een meer betrouwbare wijze beschikbaar komen.

Resultaten

Uit tabel 5.18. waarin de drie 'sociaalemotionele graadmeters' zijn opgenomen, blijkt het volgende:

1. De I/D ratio, het quotiënt van nondirectief (indirect) en directief (direct) gedrag van de leraar toont een nogal groot verschil tussen leraar B en de overige drie leraren, die niet veel van elkaar verschillen. De teller van de breuk, het totaal aantal seconden nondirectief gedrag over acht lessen, toont echter aan dat

leraar B aanzienlijk meer tijd heeft besteed aan dit gedrag dan zijn collega's en zelfs meer dan tweemaal zoveel als leraar C;

2. De ratio acceptatie/critiek (of afwijzen van gevoelens, ideeën, en dergelijke) toont wederom een groot verschil tussen leraar B en zijn collega's. Leraar A en D ontlopen elkaar hier nauwelijks, maar leraar C is hier de enige wiens critiek groter is dan het accepterend gedrag, waardoor het verschil tussen hem en leraar B wel erg groot wordt;
3. Een vergelijking ten aanzien van het percentage leerlinginitiatief dat tijdens de lessen van de diverse leraren plaatsvond, laat zien dat ook hier weinig overeenstemming bestaat tussen leraar B en leraar C. Leraar C blijkt aanzienlijk meer leerlinginitiatief toe te staan dan leraar B, respectievelijk 2,7% (39%) en 0,7% (10%). Van de twee andere leraren staat A meer initiatief toe dan D, respectievelijk 2,3% (32%) en 1,3% (19%).

	Teacher A		Teacher B		Teacher C		Teacher D	
social emotional cues	rv	results	rv	results	rv	results	rv	results
I/D-ratio	2	996/8658 = 0.12	1	1582/8593 = 0.20	4	750/9962 = 0.08	3	1041/10904 = 0.10
ratio: acceptation / criticism	2	253/96 = 2.6	1	400/58 = 6.9	4	150/158 = 0.95	3	222/93 = 2.4
pupil initiation	2	371 = 2.3%	4	115 = 0.7%	1	472 = 2.7%	3	246 = 1.3%

Tabel 5.18. : The 3 social-emotional cues: I/D-ratio, ratio acceptation (1 + 2)/criticism (12 + 13), and pupil initiation; rank order (rv) of the 4 teachers is based on the results of 8 lessons each.

In tabel 5.20. waar de frequentietotalen van belang zijn, is te zien dat leraar B 15,4% van het totaalaantal 'statements' dat hij gedurende acht lessen heeft gemaakt, heeft besteed aan accepterende en aanmoedigende opmerkingen en/of gedragingen (collectief en individueel). Leraar C komt tot 6,4%. Het verschil tussen A en D is gering, respectievelijk 10,6% en 8,5%. Ten aanzien van afwijzende opmerkingen en/of gedragingen heeft leraar B 2,3% van het totaalaantal 'statements' aan dit gedrag besteed (collectief en individueel), leraar C 6,8%. De leraren A en D ontlopen elkaar niet veel, respectievelijk 3,9% en 3,6%.

Beschouwen we dezelfde categorieën in tabel 5.19., waar de kolomtotalen de basis vormen voor de vergelijking tussen de leraren, dan komt eenzelfde beeld naar voren. De leraren B en C nemen de 'extreme' posities in, terwijl de leraren A en D zich daartussen bewegen en onderling nauwelijks verschillen.

Er zij verder nog op gewezen dat er over het algemeen bijzonder weinig vragen worden gesteld (tabel 5.19.): leraar B scoort in deze categorie het laagst, namelijk 0,6% van zijn totale lestijd; leraar D het hoogst, namelijk 2,1%.

Leraar D laat de leerling het meest voordoen, leraar A het minst, respectievelijk 2,3% en 0,9%.

Ten aanzien van het belangrijke instructiegedeelte van de les lichamelijke oefening valt op dat de leraren A en D elkaar, wat betreft de instructie aan luisterende leerlingen, nauwelijks ontlopen, maar dat de leraren B en C beiden hieraan minder tijd besteden. Precies het tegenovergestelde moet worden opgemerkt over de instructie én correctie aan leerlingen die actief bezig zijn.

Ten slotte is het van belang te constateren dat leraar B de leerlingen gedurende lange tijd opdrachten laat uitvoeren, zonder verbale tussenkomst van zijn kant.

Een voorlopige conclusie op grond van deze gegevens kan als volgt worden samengevat: op alle sociaalemotioneel relevante categorieën vormen de leraren B en C elkaars tegendeel, en ontlopen de leraren A en D elkaar niet zoveel. Het is belangrijk hierbij op te merken dat de leraren B en C les gaven aan de twee klassen van de experimentele groep. Er moet echter rekening gehouden worden met het feit

	Category	Teacher A			Teacher B			Teacher C			Teacher D			
		rv	%	(%)	rv	%	(%)	rv	%	(%)	rv	%	(%)	
1	acceptance collectively	3	0.7	(18)	1	1.7	(45)	4	0.5	(13)	2	0.9	(24)	(100%)
2	acceptance individually	2	4.1	(26)	1	6.4	(41)	4	2.4	(15)	3	2.8	(18)	"
3	joining a game	2.5	0.1	(25)	4	0.0	(0)	2.5	0.1	(25)	1	0.2	(50)	"
4	giving aid non verbal	4	0.7	(20)	1.5	1.0	(29)	1.5	1.0	(29)	3	0.8	(23)	"
5	'broad' questions	2	0.6	(32)	4	0.2	(11)	3	0.3	(16)	1	0.8	(42)	"
6	'narrow' questions	2	0.9	(26)	4	0.4	(12)	3	0.8	(24)	1	1.3	(38)	"
7	instruction inactive	1	20.9	(29)	3.5	15.2	(21)	3.5	14.9	(21)	2	19.9	(28)	"
8	showing a technique	1.5	1.6	(29)	3	1.2	(22)	1.5	1.6	(29)	4	1.0	(19)	"
9	instruction active	4	12.5	(18)	1	20.7	(30)	2	19.4	(28)	3	16.1	(23)	"
10	giving directions	2.5	14.2	(26)	4	11.4	(20)	2.5	14.3	(26)	1	15.8	(28)	"
11	explicit stimulation	3	1.0	(24)	4	0.7	(18)	1	1.2	(30)	2	1.1	(28)	"
12	criticizing collectively	2	0.9	(26)	4	0.3	(9)	1	1.5	(44)	3	0.7	(21)	"
13	criticizing individually	2.5	1.4	(22)	4	0.7	(11)	1	2.8	(44)	2.5	1.4	(22)	"
14	action and performance	2	33.6	(25)	1	38.1	(28)	3	32.4	(24)	4	31.3	(23)	"
15	showing a technique	4	0.9	(15)	2	1.7	(27)	3	1.3	(21)	1	2.3	(37)	"
16	initiative and answering 'broad' quest.	2	2.3	(32)	4	0.7	(10)	1	2.7	(39)	3	1.3	(19)	"
17	silence or confusion	1	3.9	(36)	4	1.6	(15)	2	2.9	(27)	3	2.4	(22)	"
			100%			100%			100%			100%		

Tabel 5.19. : Results of judgements of 8 lessons of physical education of four teachers each: rankorder (rv) of the teachers, *percentage (%) per category over the total time per teacher*; percentage between brackets give percentage per category between the four teachers.

Category	Teacher A			Teacher B			Teacher C			Teacher D			
	rv	%	(%)	rv	%	(%)	rv	%	(%)	rv	%	(%)	
1. acceptance collectively	3	1.8	(23)	1	3.0	(38)	4	1.1	(14)	2	2.0	(25)	(100%)
2. acceptance individually	2.	8.8	(27)	1	12.4	(38)	4	5.3	(16)	3	6.5	(19)	"
12. critizizing collectively	2	1.6	(27)	4	0.6	(10)	1	2.4	(40)	3	1.4	(23)	"
13. criticizing individually	2.5	2.3	(21)	4	1.7	(16)	1	4.4	(42)	2.5	2.2	(21)	"

Tabel 5.20. : Rank order (rv) of four teachers, percentage (%) per category over the total number of frequencies per teacher; percentage between brackets give percentage per category between the four teachers.

dat de sociaalemotioneel relevante categorieën tezamen nog geen 20% van het leraargedrag beslaan. De categorieën die iets zeggen over docerend lesgeven, namelijk instructie inactief, instructie actief, gedragsaanwijzingen en uitvoering van opdrachten door de leerling, beslaan bij alle vier leraren meer dan 80% van hun lesgeefgedrag. Dus wat betreft een zeer groot deel van hun lesgeefgedrag komen de vier leraren grofweg overeen wanneer we enkele, boven reeds genoemde, verschillen buiten beschouwing laten.

Ons standpunt is echter dat juist in de les lichamelijke oefening, die nogal collectief is opgezet, opmerkingen en/of gedragingen van de leraar, die getuigen van aandacht en 'gevoel' voor de leerling(en), de moeite van het registreren waard zijn. Op het onderscheid tussen deze twee aspecten van het leraargedrag, namelijk overeenkomst in instructie en verschil in sociaalemotionele benadering, zal in de discussie nader worden ingegaan.

5.3.3. ACTIVITEIT TIJDENS DE LESSEN LICHAMELIJKE OEFENING

Om de invloed van extra lessen lichamelijke oefening te kunnen nagaan, is het noodzakelijk om een indruk te verkrijgen van de wijze waarop de lessen door de diverse leraren worden gegeven. De reden is, dat verondersteld wordt dat twee extra lessen lichamelijke oefening per week een verhoging van het activiteitsniveau tijdens de schooluren met 2/3 zal betekenen. Hierbij werd er echter vanuit gegaan dat de lessen van alle betrokken leraren identiek zullen verlopen. Deze veronderstelling lijkt voorbarig.

Een belangrijke bijdrage tot een zo gelijk mogelijk verloop van de lessen werd reeds gevonden in het ontwerpen van een volledig lessenplan, identiek voor alle brugklassen. Desondanks is het mogelijk dat de individuele interpretatie en de persoonlijkheid van de leraren verschillen veroorzaken in de presentatie van de oefenstof. Deze verschillen kunnen zowel kwalitatief als kwantitatief van aard zijn.

Om een indruk te krijgen van de kwantitatieve verschillen in eenzelfde les, gegeven door de verschillende leraren, werd het activiteitsniveau van de leerlingen tijdens de les gemeten. Hiervoor werd gebruikgemaakt van stappentellers, op de manier zoals beschreven in 5.2..

Meetprocedure

Voor het bepalen van het activiteitsniveau tijdens de les lichamelijke oefening werd bij elke leraar een steekproef van tien lessen genomen.

Deze lessen waren voor alle vier leraren hetzelfde. Tijdens elke van deze lessen kregen vijf tevoren gekozen leerlingen een stappenteller om. De leerlingen werden allen even vaak gekozen. De stappentellers werden aan het begin van de les uitgereikt en onmiddellijk na afloop weer teruggenomen.

Resultaten

De resultaten van deze meting, weergegeven in tabel 5.21., geven aan dat de gemiddelde activiteitscores over tien lessen een verschil

van maximaal 21,7% tussen de leraren kunnen vertonen. De spreiding, ontstaan door het meten tijdens lessen met een zeer verschillend karakter, blijkt zeer groot te kunnen zijn.

De verschillen tussen de leraren ten aanzien van de activiteit van de leerlingen tijdens hun lessen werden getoetst door middel van een t-toets (De Jonge, 1963). Hieruit blijkt dat de t-waarden een significantieniveau van 5% niet bereiken.

Scores of pedometers during lessons of physical education

Teacher	A	B	C	D
mean	2880	3250	2670	3080
stand.dev.	790	570	870	920
minimum	1960	2390	1730	2150
maximum	4140	4050	3940	4460
range	2180	1660	2210	2310

Tabel 5.21. : Mean scores of pedometers during ten lessons of physical education. The same lessons had been given by four different teachers. The score of each lesson was obtained from the mean pedometer score of five randomly chosen pupils.

5.3.4. LEERLINGBEOORDELING VAN HET LERAARGEDRAG

Om na te gaan of er overeenstemming bestaat over het lesgeefgedrag van de leraar tussen beoordelingen van deskundigen betreffende sociaalemotionele en vaktechnische aspecten enerzijds, en beoordelingen door de leerlingen anderzijds, werd een vragenlijst samengesteld. Deze vraagstelling was niet voorzien in de oorspronkelijke opzet. Om toch een indruk te krijgen over de opvattingen van de leerlingen over hun leraar, werden in kort bestek een tweetal vragen ontworpen en ter beantwoording voorgelegd aan de leerlingen.

Het beoordelen van leraren door hun leerlingen wordt vaak becritiseerd. Deze bezwaren betreffen de ethische aspecten van het beoordelen, het gebrek aan ervaring van de leerlingen bij het beoordelen, en 'bias', dat de testuitkomsten kan beïnvloeden (Cogan, 1958; Remmers, 1971).

Daartegenover stellen andere onderzoekers dat effectief leren het resultaat is van de interactie tussen de leraar en de leerlingen en dat, hoewel een objectieve beoordeling vrijwel niet mogelijk is, de vragenlijsten toch waardevolle informatie kunnen verschaffen over de reacties van de leerling over het lesgeefgedrag van de leraar. Het is immers de leerling die moet leren, en zijn beeld van de leraar is van doorslaggevend belang. Bovendien ervaren de leerlingen veel meer en substantieel belangrijker lesgeefgedrag, en niet een paar korte 'samples', zoals dat het geval is wanneer bij voorbeeld hoofden van scholen of inspecteurs het lesgeefgedrag van de leraar beoordelen (Watson, 1971).

Remmers (1971) stelde vast dat, wanneer de ratings van meer dan vijfentwintig leerlingen worden gemiddeld, deze een even betrouwbaar beeld geven van de leraar als de beste 'educational and mental tests' die op dat moment bestonden.

Op het gebied van de lichamelijke oefening is echter weinig onderzoek als vergelijkingsmateriaal voorhanden, en daarom zal het aanhalen van literatuur beperkt blijven tot het noemen van enig onderzoek dat een relatie heeft met sociaalemotionele en vaktechnische aspecten van het leraargedrag. Medley e.a. (1959) vonden met hun 'My Class Inventory' dat er verband bestond tussen leerlingbeoordelingen van hun leraar en het emotionele klimaat in de klas, gemeten met behulp van een interactieanalyse.

Cogan (1958) trachtte het leraargedrag te beschrijven via de drie categorieën:

1. 'inclusief', dat wil zeggen dat de leerlingen worden betrokken bij het lesgebeuren terwijl de leraar vriendelijk en begrijpend is;
2. 'preclusief', dat wil zeggen dat de leraar autoritair optreedt en de betrokkenheid van de leerlingen met het lesgebeuren afwijst;

3. 'conjunctief', dat wil zeggen dat de leraar vaktechnisch bekwaam is.

De afhankelijke variabelen waren 'hoeveelheid uitgevoerde taken' en 'zelfwerkzaamheid', beide gemeten met een 'check list'. De onafhankelijke variabele bestond uit een meting van het leraargedrag zoals dat door de leerlingen werd beoordeeld. De meting geschiedde eveneens met behulp van een 'check list'.

Uit zijn onderzoek concludeerde Cogan via covariantieanalyse dat 'inclusief' gedrag van de 'science teacher' positief correleerde met 'zelfwerkzaamheid' en 'hoeveelheid uitgevoerde taken' bij veertien- en vijftienjarige leerlingen. Dit vond hij ook - maar in mindere mate - bij 'conjunctief' gedrag van de leraar. Hij vond geen enkele hoge correlatie met 'preclusief' gedrag. Er bestond een groot verband tussen 'inclusief' en 'conjunctief' leraargedrag.

Het instrument

Aan het einde van het cursusjaar beoordeelde elk van de vier klassen de leraar lichamelijke oefening van wie deze klassen dat jaar les hadden gehad. Onafhankelijke medewerkers namen de vragenlijst af. Tijdens de afname waren de leraren niet in de klas aanwezig. De leraren die aan de experimentele groep les gaven, zijn leraar B en leraar D. In de controlegroep gaven leraar A en leraar C les.

Het instrument bestaat uit twee vragen en beide vragen hebben drie 7-puntsschalen. Dit is weergegeven in schema 5.22..

De eerste vraag heeft betrekking op het sociaalemotionele klimaat in de klas. De tweede vraag heeft betrekking op de vaktechnische aspecten van het lesgeefgedrag van de leraar. De eerst genoemde pool van de schaal krijgt zeven punten, de andere pool krijgt één punt. De schalen 2 en 5 lopen echter van zeven punten naar één punt om eventuele 'response-set' bij het beantwoorden van de vragen te voorkomen. Hoe positiever de vraag wordt beantwoord, hoe meer punten de leraar ontvangt. De totaalscore van iedere leraar werd berekend door het gemiddelde te berekenen van de somscore van alle leerlingen in een klas per schaal.

Later bleek dat schaal 5 bij de leerlingen meer een oordeel over de gymnastiekleraar als persoon zal opwekken, en dus onder

Question 1. : 'What is your judgement of your teacher of physical education as a person?' 'Wat vind jij van de gymnastiekleraar als persoon?'			
scale 1.	pleasant aardig	(7) - - - - -(1)	not pleasant niet aardig
scale 2.	makes many remarks heeft veel aanmerkingen	(1) - - - - -(7)	makes few remarks heeft weinig aanmerkingen
scale 3.	gives you much freedom laat je vrij	(7) - - - - -(1)	gives you less freedom houdt je strak
Question 2. : 'What is your judgement of the professional skill of your teacher of physical education?' 'Wat vind je van de gymnastiekleraar als vakman, dat wil zeggen, zoals hij lesgeeft?'			
scale 4.	does it right goed	(7) - - - - -(1)	does it badly slecht
scale 5.	makes many critical remarks to the performance after instruction heeft veel critiek op de uitvoering van opdrachten	(1) - - - - -(7)	makes few critical remarks to the performance after instruction heeft weinig critiek op de uitvoering van opdrachten
scale 6.	he lets us do many nice things during the lesson doet veel leuke dingen in de les	(7) - - - - -(1)	he lets us do few nice things during the lesson doet weinig leuke dingen in de les

Schema 5.22. : Questionnaire concerning the judgement of pupils about the teaching behaviour of their teacher of physical education in two aspects:
Question 1. (scale 1., 2. and 3.) personally;
Question 2. (scale 4., 5. and 6.) professionally;
on 7-point rating scales.

vraag 1 thuishoort. Bovendien is de formulering van de schaal onjuist. Met critiek was bedoeld: negatieve critiek. Nu zou uit de vraag blijken alsof een leraar geen critiek mag uitoefenen, of geen positieve

critiek kan leveren. Om deze redenen is de vraag niet verder opgenomen in de berekeningen.

Resultaten

In tabel 5.23. zijn per klas per schaal de gemiddelden weergegeven. Tevens is de standaardafwijking van de gemiddelden en de rangvolgorde van de leraar op iedere schaal berekend.

Teacher	A			B			C			D		
Class	Gym. 1			Gym. 2			Ath/Havo 3			Ath/Havo 4		
Number of pupils	16			17			16			21		
Scale	mean	s.d.	rv	mean	s.d.	rv	mean	s.d.	rv	mean	s.d.	rv
1. pleasant	4.1	1.41	(2)	3.5	1.19	(4)	5.6	1.49	(1)	3.7	1.89	(3)
2. remarks	3.1	1.25	(2)	2.3	1.49	(4)	4.6	0.93	(1)	2.4	1.53	(3)
3. freedom	3.5	1.58	(2)	2.2	1.21	(4)	4.4	1.66	(1)	2.9	2.10	(3)
mean 1+2+3	3.6			2.7			4.9			3.0		
4. right	4.6	1.27	(4)	5.9	1.95	(1)	5.7	1.76	(2)	5.5	1.37	(3)
6. nice things	3.6	1.62	(4)	5.9	1.57	(1)	5.6	1.06	(2)	4.0	2.25	(3)
mean 4 + 6	4.1			5.9			5.7			4.8		

Tabel 5.23. : Mean, standard deviation (s.d.) and rankorder (rv) per scale of pupil judgement of their teacher.

Omdat deze vragenlijst niet eerder op bruikbaarheid werd getoetst, en omdat bovendien alleen op de natest werd gemeten, zal slechts naar de relatieve verschillen worden gekeken; wij zijn er niet toe overgegaan om de verschillen op hun significantie te toetsen. Hierdoor moet de nodige voorzichtigheid worden betracht bij het interpreteren van de gegevens.

Het hoogste gemiddelde voor vraag 1. heeft leraar C met 4,9; hij wordt daardoor het meest positief door de leerlingen beoordeeld. Het laagst wordt leraar B met 2,7 beoordeeld. De rangvolgorde van de leraren is voor alle drie schalen volledig overeenstemmend. Dit laatste geldt tevens voor vraag 2.. Het hoogste gemiddelde bij deze

vraag heeft leraar B; hij wordt dus als een goede vakman beschouwd. Het laagste gemiddelde heeft leraar A met 4,1..

De rangschikking van de leraren is zeer consistent, maar niet overeenkomstig wat betreft de twee vragen.

Literatuur

1. Amidon, E., E. Hunter, Verbal interaction in the classroom: The Verbal Interaction Category System, in: Interactionanalysis: theory, research and application, Addison/Wesley, Massachusetts, 1967, 141-149.
2. Anderson, H.H., The measurement of domination and socially integrative behavior in teacher's contacts with children, in: Interactionanalysis: theory, research and application, Addison/Wesley, Massachusetts, 1967, 4-23.
3. Bales, R.F., Interaction process analysis, a method for the study of small groups, Addison/Wesley, Massachusetts, 1950.
4. Barr, A.S., Teacher effectiveness and its correlates, in: A.S. Barr e.a., Wisconsin studies of the measurement and prediction of teacher effectiveness, A summary of investigations, Wisconsin, 1967.
5. Barr, A.S., Measurement and prediction of teacher efficiency, Rev. Educ.Res., 25, 1955, 261-269.
6. Bartlett, C.J., Dimensions of leadership behavior in classroom discussion groups, J.Educ.Psych., 50, 5, 1959, 280-285.
7. Becker, W.C., C.H. Madsen, C.R. Arnold, D.R. Thomas, The contingent use of teacher attention and praise in reducing classroom behavior problems, J. of Special Educ., I, 1967, 287-307.
8. Bergeijk, J. van, Didactisch handelen. Een terreinverkenning in verband met de practische scholing van aanstaande leraren, H. Veenman en zonen, Wageningen, 1971.
9. Block, A. de, Algemene didactiek, vooral in verband met de vorming van de leerlingen van 10 - 18 jaar, Antwerpen, 1965.
10. Bookhout, E.C. Teacher behavior in relation to the social-emotional climate of physical education classes, Res.Quart., 38,3, 1967, 336-348.
11. Calvin, A.D., F.K. Hoffman, E.L. Harden, The effect of intelligence and social atmosphere on group problem solving behavior, in: Gage, Handbook of research on teaching, Rand McNally, Chicago, 1971, 683-714.
12. Cogan, M.L., The behavior of teachers and the productive behavior of their pupils, J. of Exp.Educ., 27, 1958, 89-124.
13. Cosgrove, D.J., Diagnostic ratings of teacher performance, J. Educ.

Psych., 50, 5, 1959, 200-205.

14. Finer, A.R., A comparative study of the effects of a direct and indirect method of teaching gymnastics on the development of selected basic motor abilities, Res.Paper Phys.Educ. , 1971, 24-32.
15. Flanders, N.A., Analysing teaching behavior, Addison/Wesley, Massachusetts, 1970.
16. Haigh, G.V., W. Schmidt, The learning of subject matter in teacher-centered and group-centered classes, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1971, 683-714.
17. Hunt, G.T., School personnel and mental health, Rev. Educ.Res., 26, 1956, 502-521.
18. Jonge, H. de, Inleiding tot de medische statistiek, N.I.P.G., Leiden, 1963.
19. Lewin, K., R. Lippitt, R.K. White, Patterns of aggressive behavior in experimentally created social climates, Journ.Soc.Psych., 10, 1939, 271-299.
20. Madsen, C.H., W.C. Becker, D.R. Thomas, Rules, praise and ignoring: elements of elementary classroomcontrol, J. of Applied Behavior Analysis, I, 1968, 139-150.
21. McKeachie, W.J., Individual conformity to attitudes of classroom groups, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1971, 683-714.
22. McKeachie, W.J., Motivation, teaching methods and college learning, in: Mr.Jones (ed), Current theory and research in motivation, Un. of Nebraska, Lincoln, 1961.
23. Medley, D.M., H.E.A. Mitzel, A technique for measuring classroom behavior, J. Educ.Psych., 49, 1958, 86-92.
24. Medley, D.M., H.E.A. Mitzel, Some behavioral correlates of teacher effectiveness, J. Educ.Psych., 50, 1959, 239-246.
25. Pankratz, R., Verbal interaction patterns in the classroom of selected physics teachers, in: Interaction analysis: theory, research and application, Addison/wesley, Massachusetts, 1967, 189-209.
26. Peck, R.F., Predicting principals ratings of teacher performance from personality data, L.Educ.Psych., 50, 5, 1959, 70-75.
27. Porter, E.H., An introduction to therapeutic counseling, New York, 1965.

28. Ras H., Het gebruik van een analyse model bij de observatie en evaluatie van lesgeefgedrag, De Lich.Opv., 23, 1971, 962-967.
29. Remmers, H.H., Rating methods in research on teaching, in: N.L. Gage, Handbook of research in teaching, Rand McNally, Chicago, 1971, 329-378.
30. Ryans, D.G., E. Wandt, A factor analysis of observed teacher behavior in the secondary school, a study of criterion data, Educ.Psych. Measurement, 12, 1952, 574-586.
31. Ryans, D.G., A study of criterion data (a factor analysis of teacher behavior in elementary school), Educ.Psych.Measurement, 12, 1952, 333-344.
32. Ryans, D.G., Notes on the criterion problem in research, with special reference to the study of teacher characteristics, J. Genet. Psych., 91, 1957, 33-61.
33. Simon, A. , E.G. Boyer, Mirrors for behavior. An anthology of classroom observation instrument,in: Bergeijk, J. van, Didactisch handelen, Wolters Noordhoff N.V., Groningen, 1971.
34. Solomon, D., W. Bezdek, L. Rosenberg, Dimensions of teacher behavior, J. of Exp.Educ., (received Febr. 26), 1963.
35. Stellwag, H.W.F., R. Vuyk, P.C. van de Griend, De leraar en zijn klas, Groningen, 1963.
36. Symonds, P.M., Characteristics of the effective teacher based on pupil evaluations, J.Exp.Educ., 23, 1955, 289-310.
37. Tausch, R., A. Tausch, Psychologie van opvoeding en onderwijs, Van Loghem Slaterus, Arnhem, 1967.
38. Veen, P., Meebeslissen, een veldexperiment in een hockeyclub, Proefschrift Rijksuniversiteit Utrecht, Van Gorcum en comp. N.V., Assen, 1969.
39. Watson, F.G., Research on teaching science, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1971, 1031-1059.
40. Wispé, L.G., Evaluating section teaching methods in the introductory course, in: N.L.Gage, Handb. of research on teaching, Rand McNally, Chicago, 1971, 683-714.
41. Withall, J.G., The development of a technique for the measurement of social-emotional climate in classrooms, J. of Exp.Educ., 17, 1949, 347-361.

hoofdstuk 6

STATISTISCHE METHODEN

- Selectie van proefpersonen
- Indeling in groepen
- Probleemstelling

6.1. BETROUWBAARHEID

6.1.1. REPRODUCEERBAARHEID ANTROPOMETRISCHE VARIABELEN

6.1.2. REPRODUCEERBAARHEID VAN DE BEOORDELINGEN VAN DE VAKTECHNISCHE ASPECTEN VAN HET LESGEEFGEDRAG

6.1.3. BETROUWBAARHEID VAN DE BEOORDELINGEN VAN DE SOCIAALEMOTIONELE ASPECTEN VAN HET LESGEEFGEDRAG

- reproduceerbaarheid

. objectiviteit
. rangvolgorde der leraren

6.1.4. REPRODUCEERBAARHEID SKELETLEEFTIJD

6.1.5. BETROUWBAARHEID STAPPENTELLERS

6.2. FACTORANALYSE

. reductie per groep variabelen
. groepen van variabelen
. niet in de groepen opgenomen variabelen
. statistische consequenties van verschillen in factorstructuur

6.3. COVARIANTIEANALYSE

6.4. NIET GEREDUCEERD TWEESTEEKPROEVEN PROBLEEM

6.5. VERGELIJKING VAN DE UITGEVOERDE ANALYSE MET EEN ANDERE METHODE

6.6. ANALYSE VAN GYMNOLOGISCHE VARIABELEN

LITERATUUR

6. STATISTISCHE METHODEN

Selectie van proefpersonen

Als proefpersonen fungeerden mannelijke leerlingen van brugklassen van een school voor VWO en HAVO in Amsterdam. Vanwege de gecompliceerde en veeleisende opzet van het onderzoek is de keuze gevallen op het St. Ignatius College, mede op grond van de grote bereidheid van bestuur, directie en leraren om aan het onderzoek mee te werken. Van een randomselectie van proefpersonen uit 12 en 13-jarige Amsterdamse scholieren kan niet worden gesproken (zie 2.).

Indeling in groepen

De 70 proefpersonen werden door een toevalsmechanisme ingedeeld (alfabetisch) in twee groepen, elk bestaande uit twee parallelklassen. Door loting werd bepaald welke van de twee parallelklassen aan de experimentele groep, en welke aan de controlegroep werd toegewezen. De experimentele groep bevatte 33 proefpersonen, de controlegroep 37.

Aan elke proefpersoon werd gemeten een reeks van 77 variabelen $X = (X_1,\ldots,X_{77})$ aan het begin van de cursus (de voortest) en een reeks variabelen $X' = (X'_1,\ldots,X'_{77})$ aan het eind van de cursus (de natest). Onderzocht wordt de reeks van verschilscores, $Z = X' - X = (Z_1,\ldots,Z_{77})$. Hoofdletters in de notatie duiden erop dat de betreffende variabelen stochastische variabelen zijn.

Probleemstelling

Het onderzoek heeft ten doel door middel van statistische toetsing van hypothesen na te gaan of er systematische verschillen in de gemiddelden van de verschilscores tussen de experimentele en de controlegroep bestaan. Bij statistisch onderzoek formuleert men, wanneer geen voorspellingen uit voorafgaand onderzoek of literatuurgegevens kunnen worden gedaan, een hypothese, nulhypothese genaamd, op dusdanige wijze dat verwerping van deze hypothese leidt tot bevestiging van een door de onderzoeker gekoesterd vermoeden, alternatieve hypothese genaamd. In dit geval kan de nulhypothese geformuleerd worden als

$$H_0 : E\,Z_i^{(e)} = E\,Z_i^{(c)} \qquad i = 1,\ldots,77.$$

Dat wil zeggen dat de verwachting E (van Expectation) van de

verschillen voor experimentele en controlegroep gelijk zijn, terwijl de alternatieve hypothese behelst dat deze verwachtingen verschillen in een voor elke i opgegeven richting (De bovenindices $^{(e)}$ en $^{(c)}$ duiden op de experimentele en controlegroep).

Dit is een multivariaat tweesteekproevenprobleem, dat onder enige niet al te restrictieve voorwaarden zeer wel statistisch kan worden getoetst. Om verschillende redenen is dit echter niet direct uitgevoerd.

In de eerste plaats werd het mogelijk geacht om de variabelen op theoretische, dat wil zeggen niet-statistische gronden in te delen in een aantal groepen van één of meer variabelen, zodat deze groepen conceptueel onafhankelijk zijn.

Ten tweede kan worden opgemerkt dat om redenen van efficiëntie, interpreteerbaarheid van de resultaten en overzichtelijkheid een drastische reductie van het aantal variabelen, die dikwijls onderling afhankelijk zijn, geboden is. De ratio variabelen/proefpersonen bedroeg 77/70.

Ten slotte bestond de behoefte om rekening te houden met de storende invloed van enige externe variabelen, de zogenaamde concomitanten, eerder aangeduid als storende variabelen (zie 5.).

Op grond van de eerste overweging werd elk van deze groepen van variabelen als een op zich zelf staand probleem behandeld.

6.1. BETROUWBAARHEID

Een instrument meet des te nauwkeuriger naarmate één meetuitkomst gemiddeld meer relevante informatie verstrekt met betrekking tot de waarde van de bijbehorende variabele (De Groot, 1968).

Van een aantal metingen hebben wij de meetbetrouwbaarheid gebaseerd op een herhaling van de metingsprocedure zonder dat de te meten objecten intussen zelf veranderd of door de eerste meting beïnvloed zijn.

Vervolgens hebben wij tussen de twee reeksen van uitkomsten (eerste meting, tweede meting) de correlatie berekend. Deze meetbetrouwbaarheidscoëfficiënt zal dichter bij +1 liggen naarmate het relatieve effect van toevalsfluctuaties geringer is.

Er zijn twee betrouwbaarheidsaspecten te onderscheiden:

1. reproduceerbaarheid;
2. objectiviteit.

De reproduceerbaarheid van een test heeft betrekking op de meetbetrouwbaarheid als deze herhaald wordt door dezelfde proefleider bij dezelfde proefpersonen en onder dezelfde omstandigheden.

De objectiviteit van een test heeft betrekking op de meetbetrouwbaarheid als deze test herhaald wordt bij dezelfde proefpersonen onder dezelfde omstandigheden, maar met een verschillende proefleider. De objectiviteit heeft dus betrekking op de invloed van de proefleider op de meetresultaten (Barrow e.a., 1968). De 'mens' blijkt een onbetrouwbaar instrument te zijn. Door standaardisering van alle instructies der meetprocedures, kan ernaar gestreefd worden dat de invloed van de proefleider op het meetresultaat zo gering mogelijk wordt. Bij overigens zeer eenvoudige metingen, zoals bij voorbeeld het meten van antropometrische variabelen, kunnen desondanks systematische verschillen optreden tussen proefleiders (Kemper e.a., in press). In het algemeen is dan ook de critiek juist ten aanzien van de tekorten in objectiviteit.

In dit onderzoek is er zorg voor gedragen dat dezelfde proefleider dezelfde variabelen heeft gemeten in controlegroep en experimentele groep. Omdat elke variabele in principe tweemaal, zowel op de voortest als op de natest, werd gemeten, is bovendien elke variabele op beide meettijdstippen door dezelfde proefleider uitgevoerd.

Een gebrek aan objectiviteit hoeft niet alleen op te treden door het ontbreken van een expliciete instructie, het kan ook zijn dat de instructies zelf een beroep doen op een beoordeling of waardering, waarvoor geen volledig objectieve normen zijn aangegeven. Bij de ontwikkeling van de twee beoordelingssystemen voor het lesgeefgedrag van de leraar lichamelijke oefening (zie 5.3.1. en 5.3.2.) werd gebruikgemaakt van een vijftal beoordelaars. In dit hoofdstuk zal worden nagegaan in hoeverre deze beoordelaars tot overeenstemming kwamen voor wat betreft hun oordeel over het lesgeefgedrag van de leraren.

Het ideaal van absolute objectiviteit zal daarbij natuurlijk

niet worden bereikt, maar er is naar gestreefd dit zo dicht mogelijk te benaderen (De Groot, 1968).

6.1.1. REPRODUCEERBAARHEID ANTROPOMETRISCHE VARIABELEN

De antropometrische metingen in dit onderzoek werden uitgevoerd door twee getrainde personen. Ieder van hen was verantwoordelijk voor de meting van een aantal tevoren bepaalde variabelen. De metingen werden op zowel de voor- als natest door dezelfde persoon uitgevoerd.

Voorafgaande aan het onderzoek werd nagegaan in welke mate deze metingen reproduceerbaar waren. Daartoe werden twintig proefpersonen tweemaal door dezelfde onderzoeker gemeten. De reproduceerbaarheidscoëfficiënten van alle antropometrische metingen zijn weergegeven in tabel 6.1..

Reproducibility of anthropometric variables		
measurement	$\bar{d}$	r
height	0.00 cm	.9997
biacromial diameter	-0.19 cm	.9790
biiliocristal diameter	+0.09 cm	.9872
upperarm circumference	-0.09 cm	.9735
thigh circumference	-0.28 cm	.9905
calf circumference	-0.02 cm	.9944
wrist breadth	+0.05 cm	.9439
bicondylar femur	+0.01 cm	.9755
bicipital skinfold	+0.04 mm	.9356
tricipital skinfold	+0.02 mm	.9752
subscapular skinfold	-0.04 mm	.9553
suprailiacal skinfold	+0.10 mm	.9038

Tabel 6.1.: Test-retest correlations of anthropometric variables ($n = 20$);
$\bar{d}$ = mean differences, r = Pearson correlation coefficients.

Uit de gegevens blijkt dat de Pearson correlatiecoëfficiënt varieert van minimaal 0,9038 (huidplooi suprailiaca) tot maximaal 0,9997 (lengte). Uit de gemiddelde verschilscores blijkt dat bij geen enkele

variabele grote systematische verschillen aanwezig zijn.

6.1.2. REPRODUCEERBAARHEID VAN DE BEOORDELINGEN VAN DE VAKTECHNISCHE ASPECTEN VAN HET LESGEEFGEDRAG

Om de reproduceerbaarheid van de beoordelingen op de vaktechnische aspecten van het lesgeefgedrag vast te stellen, is één van de vier lessen voor een tweede maal ter beoordeling aangeboden aan de vijf beoordelaars. Van te voren werd niet aan hen meegedeeld dat het ging om een tweede beoordeling van deze les. Men bemerkte overigens echter vrij snel dat het ging om een reeds eerder vertoonde les (zie 5.3.1.).

De les die voor de tweede maal werd beoordeeld, was les 6, waarin kopstand werd beoefend. Deze les werd gegeven in de periode tussen Kerstmis en Pasen. Het interval tussen de eerste en de tweede beoordeling bedroeg zes maanden.

In verband met het gebruik van een 5-puntsschaal konden de verschillen tussen beide beoordelingen niet groter zijn dan 4 punten, namelijk van 5 = uitstekend naar 1 = slecht, en omgekeerd.

De verschilscores van de vijf beoordelaars afzonderlijk per leraar bleken een waarde te kunnen hebben van -3, -2, -1, 0, +1, of +2 (zie tabel 6.2.). Volgens deze indeling werden de verschilscores geturfd en gesommeerd over de tien categorieën; een - of + betekent dat de tweede beoordeling respectievelijk lager of hoger is dan de eerste beoordeling.

Uit de resultaten blijkt dat het aantal gelijkblijvende scores (verschilscore = 0) per beoordelaar en ook per docent verschillend is. Ten aanzien van de verschilscores groter dan 0 valt op dat de meeste negatief zijn, dat wil zeggen, dat tijdens de tweede beoordeling de leraren als minder goed werden beoordeeld.

Van bijzonder belang zijn echter de verschilscores ter grootte van nul. Hoe groter dit aantal, hoe hoger de mate van betrouwbaarheid. Wanneer wij ten aanzien van het aantal gelijkblijvende beoordelingen een percentage van 70% redelijk achten, kunnen wij, uitgaande van een betrouwbaarheidsniveau van 5%, linkseenzijdig de betrouwbaarheidsgrens berekenen. Het bleek nu dat waarden van 17 en hoger acceptabel genoemd konden worden. Op grond hiervan komen wij tot de

conclusie dat de beoordelaars 1, 2 en 3 ieder twee van de viermaal deze waarde niet bereiken, beoordelaar 5 slechts eenmaal niet, terwijl beoordelaar 4 geen enkele keer hieraan weet te voldoen. Over het totaal wordt dan gevonden dat elf van de twintig keer niet aan de door ons gestelde norm wordt voldaan (zie tabel 6.2.).

Van het totale aantal verschilscores groter dan nul nemen de negatieve 64% voor hun rekening. Er komen zelfs veranderingen voor van 3 punten. Beoordelaar 1 bijvoorbeeld gaf een waardering van 4 (= goed) op de eerste beoordeling en een 1 (= slecht) op de tweede beoordeling; veranderingen van 2 punten traden 32 en 3 maal op, dat is 12% van het totale aantal verschilscores groter dan nul.

		difference-scores					
judges	teachers	-3	-2	-1	0	+1	+2
1	A		3	4	17	6	
	B	1	5	11	11*	2	
	C			4	18	6	2
	D	1	3	14	11*		1
2	A		1	15	14*		
	B		2	8	18	2	
	C		1	13	14*	2	
	D		1	2	19	8	
3	A		4	22	4*		
	B		1	14	12*	3	
	C		1	5	24		
	D		1	5	18	6	
4	A		3	17	8*	2	
	B		3	9	15*	3	
	C		2	16	12*		
	D			11	16*	3	
5	A		1	7	18	4	
	B			2	22	6	
	C			4	17	9	
	D			6	16*	8	
Total		2	32	189	304	70	3

Tabel 6.2. :- The difference-scores of the 5 judges between the first and the second judgement, classified in magnitude of the difference and summed over the 10 subcategories;
- or + means respectively: lower or higher on the second judgement ($^* = p \leq .05$).

Ook kan de vraag gesteld worden of bepaalde categorieën beter of slechter worden beoordeeld dan andere categorieën.

Daarvoor zijn van alle beoordelaars tezamen de beoordelingen ingedeeld per categorie volgens hetzelfde principe als in tabel 6.2.. Als wij ook nu weer uitgaan van een percentage van 70%, wordt de linker betrouwbaarheidsgrens gevormd door de waarde 35 bij een onbetrouwbaarbeidsniveau van 5%.

Uit tabel 6.3. blijkt dan dat geen enkele categorie deze waarde bereikt. Het is dus niet mogelijk een of meer categorieën aan te wijzen, waarin de beoordelingen volgens de door ons gestelde normen betrouwbaar genoemd kunnen worden.

	d i f f e r e n c e - s c o r e s					
categories	-3	-2	-1	0	+1	+2
1.1. adjustment	1	4	13	29*	10	3
1.2. methods		8	17	31*	4	
1.3. corrections		4	20	29*	7	
2.1. formation		7	22	29*	2	
2.2. safety	1	3	24	31*	1	
3.1. working intensity			23	32*	5	
3.2. joy in working		2	22	29*	7	
4.1. attitude			16	33*	11	
4.2. speech			11	34*	15	
4.3. explanation and demonstration		4	20	28*	8	
Total	2	32	189	304	70	3

Tabel 6.3. : The difference-scores per category, summed over all judgements of the judges, and classified in magnitude of the difference;
- or + means respectively: lower or higher on the second judgement ($^{*} = p \leq .05$).

6.1.3. BETROUWBAARHEID VAN DE BEOORDELINGEN VAN DE SOCIAALEMOTIONELE ASPECTEN VAN HET LESGEEFGEDRAG

Zoals onder 5.3.2. reeds vermeld, werd voor de analyse van het lesgeefgedrag van de leraren lichamelijke oefening gebruikgemaakt van een interactieanalyseschema, gebaseerd op Flanders. Vijf beoordelaars beoordeelden, onafhankelijk van elkaar, op videotape geregistreerde lessen. Het aantal van vijf is gekozen om een zo betrouwbaar mogelijk beeld te verkrijgen van het lesgeefgedrag van de vier leraren en hun onderlinge positie op een aantal aspecten van hun onderwijsstijl. Bovendien is het op deze wijze mogelijk zogenaamde extreme beoordelaars te elimineren.

Reproduceerbaarheid

Voor de bepaling van de mate van overeenkomst tussen de meetresultaten op de 17 categorieën over twee waarnemingen van één beoordelaar is gekozen voor een methode van Scott (1955). Deze methode is relatief ongevoelig voor lage frequenties, verder kan zij worden toegepast op fracties of procenten. De methode is ontwikkeld als een index voor 'intercoder-agreement', dat wil zeggen voor een bepaling van de mate van overeenstemming tussen meetresultaten van verschillende beoordelaars. In afwijking daarvan wordt de methode hier gebruikt als een maat, in Scott's terminologie voor 'intracoder-agreement'. De essentie van Scott's methode en de gemodificeerde versie ervan die door Flanders (1960, 1967) is beschreven, wordt hierdoor volgens ons niet aangetast. In het kort komt deze hierop neer:

1. het totaal per categorie wordt omgerekend in een percentage van het totaal over alle 17 categorieën;
2. het verschil tussen de aldus berekende percentages per categorie voor de twee waarnemingen wordt gesommeerd. Op deze wijze vindt men het totaalpercentage verschil (tussen beide waarnemingen);
3. dit percentage 'werkelijk' verschil wordt in een formule van Flanders (1967) vergeleken met de overeenstemming tussen beide waarnemingen, die op grond van kansberekening te verwachten is. Π , zoals Scott zijn coëfficiënt noemt, is dan de mate waarin de betrouwbaarheid het toeval overschrijdt. De coëfficiënt houdt rekening met het aantal categorieën van het schema en de frequentie per categorie.

Een volledige beschrijving van de gemodificeerde Scott-index, zoals

deze in dit onderzoek is gebruikt, is te vinden bij Flanders (1960, 1967).

judge	interval of 3 months		interval of 2 months		interval of 1 month		interval of 3 months	
	lesson 4	teacher A	lesson 5	teacher D	lesson 7	teacher B	lesson 8	teacher D
B2	- -		.74		.74		.91	
B3	.81		.77		.91		.67	
B4	.85		.75		.83		.90	

Tabel 6.4. : Agreement between two observations of one judge, calculated with Scott's π .

Door organisatorische problemen konden helaas geen betrouwbaarheden worden berekend van de beoordelaars B1 en B5. Om dezelfde reden viel een betrouwbaarheidsbepaling van beoordelaar B2 weg. Voor het vaststellen van de reproduceerbaarheid der metingen is de volgende procedure gehanteerd: de beoordelaars beoordeelden, na verschillende intervallen, 'at random' gekozen lessen van de verschillende leraren voor een tweede maal, waarbij ernaar gestreefd werd het interval tussen eerste en tweede waarneming niet groter dan drie maanden te maken (zie tabel 6.4.). Flanders (1960, 1967) noemt een (gemodificeerde) Scott-coëfficiënt van 0,85 redelijk, overigens volkomen willekeurig, zonder nadere argumentatie. De gemiddelde betrouwbaarheidscoëfficiënten van de beoordelaars liggen, zoals tabel 6.4. laat zien, rondom 0,80. Hoewel deze waarden zeker niet optimaal genoemd kunnen worden, menen wij toch dat ze voldoende zijn om de beoordelaars als redelijk intern consistent te kunnen kwalificeren, mede gelet op de grote intervallen tussen eerste en tweede waarneming. Ook wanneer men middelt over de kolomtotalen blijkt de reproduceerbaarheid van het instrument aan redelijke normen te voldoen.

Objectiviteit

Het waarnemingsmateriaal waarover men de objectiviteit van het meetinstrument, hier geoperationaliseerd als de onderlinge overeenstemming

tussen de beoordelaars, moet bepalen, laat zich in eerste instantie karakteriseren in schema 6.5..

8 lessons of physical education given by teacher X																		
Judges $i = 1,\ldots,5$	categories ($j = 1,\ldots,7$)																	row-totals $\sum_{j=1}^{17} f_{ij} = 1$
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
B1																		1
B2																		1
B3								f_{ij}										1
B4																		1
B5																		1
	column totals																	

Schema 6.5. : Survey of the observation materials of 5 judges concerning 17 categories of teacher X.

In schema 6.5. vormt f_{ij} de fractie tijd die volgens beoordelaar i door leraar X in categorie j wordt doorgebracht. Men moet zich voorstellen dat de cellen gevuld zijn met het, per beoordelaar, over acht lessen gemiddelde (of gesommeerde) aantal waarnemingen, omgezet in fracties van de totaaltijd, per categorie. Omdat er vier leraren zijn, zouden er in principe vier van dergelijke schema's nodig zijn. Omdat de beoordelaars waarnemingen hebben gedaan over dezelfde leraar en dezelfde lessen, en omdat er met fracties wordt gewerkt, is gegeven dat $\sum_{j=1}^{17} f_{ij} = 1$ voor alle $i = 1,\ldots,5$. We wensen de nulhypothese te toetsen: H_0: 'geen verschillen in score tussen de beoordelaars'. Indien H_0 wordt verworpen, willen we bovendien nog de 'afwijkende' beoordelaars aanwijzen.

Gebleken is dat de statistische analyse ernstig wordt bemoeilijkt door de afhankelijkheid (onderlinge correlatie) van (tussen) de f_{ij}'s op één rij, omdat de rijtotalen vast zijn (= gelijk aan 1). Een χ^2-toets (als toets op homogeniteit) op fracties in een k x r-tabel met vaste rijtotalen kan niet worden toegepast omdat ten eerste de grootheden f_{ij} niet als binomiale fracties in de zin van fractie

successen bij n onafhankelijke experimenten kunnen worden opgevat en omdat ten tweede de (verwachte) aantallen in de cellen soms erg klein zijn, zodat de χ^2-approximatie faalt.

Vanwege de gesignaleerde afhankelijkheid van de f_{ij}'s op één rij, en omdat normaliteitsveronderstellingen omtrent dit waarnemingsmateriaal niet valide zijn, is uiteindelijk gekozen voor Kendall's concordantiecoëfficiënt W, een maat ter bepaling van de relatie tussen verscheidene rangordeningen van n objecten of personen. Deze maat is zeer geschikt voor de bepaling van een zogenaamde 'interjudge-reliability' (Siegel, 1956).

De volgende procedure is gevolgd:
Per categorie (om de onderlinge afhankelijkheid te vermijden) en per beoordelaar is het totaal aantal seconden of het totaal aantal frequenties over acht lessen gerangordend over de vier leraren, waarover de waarnemingen zijn gedaan. Met behulp van Kendall's W verkrijgt men aldus een indruk van de overeenstemming tussen de beoordelaars. Een meer uitgebreide beschrijving van deze concordantiemaat is hier niet opgenomen vanwege de grote bekendheid van deze techniek. Conover (1971) en vooral Siegel (1956) bespreken deze techniek uitgebreid en duidelijk.

Een belangrijk nadeel van Kendall's W is dat de bepaling van de intersubjectieve overeenstemming ongevoelig is voor absolute waarden en verschillen daartussen. Echter, gezien het uitgangspunt dat verschillen in onderwijsstijl tussen de leraren op exploratieve wijze zullen worden onderzocht (zie 5.3.2.), is het verkrijgen van een betrouwbare rangordening van de leraren op relevante aspecten van hun lesgeefgedrag een minimale voorwaarde waaraan moet worden voldaan, om althans van objectiviteit te kunnen spreken. Aan deze voorwaarde voldoet Kendall's W : hij doet een uitspraak over de onderlinge overeenstemming, in rangordening van de leraren tussen de vijf beoordelaars.

Kendall's concordantiecoëfficiënt is niet afhankelijk van allerlei veronderstellingen omtrent de onderliggende verdeling van de gegevens. Bovendien heeft hij het voordeel - in vergelijking bij voorbeeld met Friedman's 'two-way analysis of variance' (Siegel, 1956)over niet-parametrische gegevens - dat de overeenstemming wordt uitgedrukt

in een cijfer tussen 0 en 1.

Categories		Column totals	Frequency totals
1	acceptance collectively	1.00**	1.00**
2	acceptance individually	.86**	.94**
3	joining a game	.30	
4	giving aid	.52*	
5	'broad' questions	.61*	
6	'narrow' questions	.90**	
7	instruction inactive	.79**	
8	showing a technique	.58*	
9	instruction active	.94**	
10	giving directions	.86**	
11	explicit stimulation	.26	
12	criticizing collectively	.85**	.81**
13	criticizing individually	.90**	.71**
14	action and performance	.75**	
15	showing a technique	1.00**	
16	initiative and answering 'broad' questions	.90**	
17	silence or confusion	.78**	

Tabel 6.6. : Kendall's W per category of the interaction analysis; (* = $p \leq .05$ ** = $p \leq .01$)

Uit tabel 6.6. blijkt dat over het algemeen sprake is van zeer grote overeenstemming tussen de beoordelaars wat betreft de rangordening van de leraren over de aspecten van hun lesgeefgedrag, die in de categorieën worden omschreven.

Slechts over twee categorieën, meespelen en expliciete stimulatie, bestaat geen overeenstemming, een resultaat dat waarschijnlijk wordt veroorzaakt door het zeer geringe aantal waarnemingen binnen deze

categorieën die verder betrekkelijk onbelangrijk bleken te zijn. De vier categorieën waarover de frequentietotalen zijn vergeleken, zijn van belang in verband met de berekening van een belangrijk aspect van het leraargedrag (zie 5.3.2.).

Een hoge (significante) concordantiecoëfficiënt betekent natuurlijk niet dat de rangordening correct is. In dit geval geldt de rangvolgorde der leraren, op grond van onafhankelijke waarnemingen van beoordelaars, voor ons als een min of meer objectieve norm omdat een - relevant extern criterium niet voorhanden is.

Rangvolgorde der leraren

Met behulp van de concordantiemaat van Kendall komt men op een zeer groot aantal van de categorieën tot een relatief objectieve rangordening van de leraren op relevante aspecten van het lesgeefgedrag. Vanwege het exploratieve karakter van dit onderdeel van het onderzoek is een meer uitgebreide statistische analyse van deze rangvolgorde achterwege gelaten. Met name zou men hierbij kunnen denken aan uitspraken over de grootte en betekenis van absolute verschillen (per categorie) tussen de vier leraren. Op grond van de gehanteerde methode lijkt een uitspraak over deze absolute verschillen minder gerechtvaardigd dan een redelijk betrouwbare schatting van de onderlinge positie der leraren op relevante gedragsaspecten, hetgeen met behulp van Kendall's W in voldoende mate wordt bereikt.

6.1.4. REPRODUCEERBAARHEID SKELETLEEFTIJD

De beoordeling van de biologische leeftijd vanaf röntgenfoto's van het hand-pols gewricht, volgens de methode van Tanner-Whitehouse-Healy, werd uitgevoerd door een speciaal hiervoor getraind persoon.

De reproduceerbaarheid werd nagegaan aan de hand van een tweemalige beoordeling van twintig röntgenfoto's. Zowel de foto's van de voor- als natest werden tegelijkertijd na afloop van het onderzoek beoordeeld. Tijdens deze procedure werden van de reeds gescoorde foto's twintig stuks toegevoegd aan de stapel nog niet gescoorde foto's, zonder de beoordelaar hiervan op de hoogte te brengen. Deze dubbele beoordeling werd gebruikt als reproduceerbaarheidsmeting.

De correlatie tussen de eerste en de tweede beoordeling van de skeletleeftijd bedroeg 0,8435. Deze waarde komt goed overeen met de correlatiecoëfficiënt die Medicus e.a. (1971) verkrijgen (r = 0,83), maar is beduidend lager dan bij Beunen (1970) die tot 0,99 komt.

De gemiddelde verschilscore tussen de eerste en tweede beoordeling bedroeg + 0,135 jaar.

6.1.5. BETROUWBAARHEID STAPPENTELLERS

De stappentellers, instrumentjes die verticale bewegingen registreren en sommeren, en die werden toegepast voor het meten van de buitenschoolse lichamelijke activiteit en van de activiteit tijdens de lessen lichamelijke oefening, werden, voorafgaande aan het onderzoek, getest op hun betrouwbaarheid. Daartoe werd een horizontale beugel, waaraan tegelijkertijd veertien stappentellers bevestigd waren, mechanisch op en neer bewogen. De bewegingen werden opgewekt door een decoupeerzaag waaraan de beugel was bevestigd. Op deze wijze was een identieke verplaatsing van alle aan de beugel bevestigde stappentellers verzekerd.

De tellers werden onderling verbonden om ook slingerbewegingen aan het beweegbare ophangpunt zoveel mogelijk identiek te houden.

De betrouwbaarheid werd bepaald door de scores te vergelijken van de stappentellers die tegelijkertijd waren bevestigd. De test werd uitgevoerd over perioden varierend van vijftig minuten tot zeseneenhalf uur.

Uit deze vergelijking blijkt dat, onafhankelijk van de lengte van de periode waarover de meting werd uitgevoerd, het verschil tussen de minimum en de maximum score slechts 0,04 tot 1,1% van de gemiddelde score bedroeg.

Ondanks deze goede resultaten zijn enige kanttekeningen van negatieve aard op zijn plaats.

De drie wijzers, behorende bij de schaal voor de honderdtallen, duizendtallen en tienduizendtallen, bleken onder invloed van het gebruik enigszins te kunnen verschuiven. Aangezien de schalen, met uitzondering van de schaal 0 - 100, zeer klein zijn, betekent een lichte verschuiving een mogelijk flinke afleesfout. Hoewel door vergelijking van de stand van de wijzers der diverse schalen eventuele

afleesfouten voor het grootste deel tijdig gecorrigeerd konden worden, blijft de mogelijkheid van incidentele fouten als gevolg van grotere verschuivingen van de wijzers mogelijk.

Daarnaast bleek bij proefnemingen met verscheidene stappentellers, bevestigd op verschillende plaatsen aan de binnenzijde (schommelvrij) van de broekband van één individu, dat de scores enigszins beïnvloed werden door de plaats van bevestiging. Aangezien de tellers aan de voorzijde van de broek de meest overeenkomende resultaten opleverde, is deze plaats verkozen voor de bevestiging van de stappentellers in dit onderzoek. De proefpersonen werd nadrukkelijk verzocht, bij het veranderen van broek de stappenteller direct links of rechts van de sluiting aan de binnenzijde te bevestigen.

6.2. FACTORANALYSE

Reductie per groep variabelen

Gepoogd werd om door middel van een factoranalyse op de voortestgegevens het aantal variabelen in een groep tot één à twee te reduceren zonder veel informatie te verliezen. De factoranalyses werden uitgevoerd op de voortestgegevens om een eventuele invloed van de onafhankelijke variabele op de afhankelijke variabelen te vermijden. Factoranalyse is een techniek om een beperkt aantal variabelen te selecteren, zodanig dat alle variabelen bij benadering als lineaire combinaties van de geselecteerde zijn te beschouwen. Men zegt wel dat de geselecteerde variabelen de gehele groep van variabelen representeren. Wil men de vergelijking van twee groepen proefpersonen op deze representanten baseren, dan is het een dwingende eis dat in beide groepen dezelfde factorstructuur kan worden gevonden. Dit vereist een gelijke correlatiestructuur. Wordt een dergelijke overeenkomst niet aangetroffen, dan is het niet mogelijk om variabelen te selecteren die voor beide groepen gelijkelijk als representant kunnen dienen. De factoranalyses werden uitgevoerd volgens de methode van Jöreskog (1963), met varimaxrotatie, voor experimentele en controlegroep afzonderlijk. Als correlatiemaat werd Spearman's rangcorrelatiecoëfficiënt gekozen (Kendall, 1962), omdat deze maat gevoelig is voor velerlei soorten samenhang.

Naast een berekening van correlatiematrices voor de conceptuele groepen afzonderlijk, werden ook correlatiematrices voor alle variabelen tezamen berekend. Dit werd gedaan om na te gaan of variabelen uit verschillende conceptuele groepen inderdaad laag correleerden.

Op grond van de toevalsindeling van de proefpersonen over experimentele en controlegroep zou men een grote overeenkomst tussen beide groepen op de voortest kunnen verwachten. Niettemin is het zinvol om deze overeenkomst aan een nadere inspectie te onderwerpen, dit om na te gaan of bij voorbaat niet al te grote verschillen bestaan die de voortest-natest opzet in gevaar brengen (Oud e.a., 1971).

In zeer veel sociaalwetenschappelijke publicaties wordt een voortest-natest opzet gehanteerd, waarin gebruikgemaakt wordt van t-toetsen voor de toetsing van de gelijkheid der proefgroepen op de voortest. In dergelijke typen onderzoek toetst men de nulhypothese: 'de proefgroepen zijn aanvankelijk gemiddeld gelijk op een aantal variabelen'. De in dit onderzoek gekozen benadering stelt bovendien de vraag: 'Is de onderlinge samenhang van variabelen, te specificeren in factoren, in beide proefgroepen dezelfde?'

In groepen variabelen waar bovengenoemde overeenkomst in factorstructuur tussen experimentele en controlegroep niet behoeft te worden verworpen, worden ten slotte representanten gekozen. Tevens wordt nagegaan of de gekozen representanten niet een te grote samenhang vertonen. Voor deze variabelen wordt een covariantieanalysemodel (zie 6.3.) opgesteld en geanalyseerd.

In groepen waar daarentegen verrassenderwijs geen voor experimentele en controlegroep gemeenschappelijke factorstructuur wordt aangetroffen, alsmede bij dié groep variabelen, waar geen factoranalyse wordt uitgevoerd omdat geen zinvolle factorstructuur aanwezig wordt geacht, kan een dergelijke analyse geen doorgang vinden. Hier is geen mogelijkheid tot verantwoorde reductie en in deze gevallen wordt het multivariaat tweesteekproevenprobleem ongereduceerd opgelost.

Groepen van variabelen

Het grootste deel van de 77 variabelen werd in acht min of meer conceptueel onafhankelijke groepen ingedeeld. Voor elk van deze groepen

is onderzocht of het mogelijk was de variabelen per groep samen te vatten in een of meer variabelen die volgens de factoranalyse hoog laden op die factor. Deze variabelen kunnen dan als min of meer onafhankelijke representanten worden opgevat. De groepen zijn steeds zo ingedeeld dat de eventuele representanten zinvol zijn en interpretatiemogelijkheden bieden. Ten aanzien van de groepen I en II is een deel van de argumentatie terug te vinden in hoofdstuk 4.1.. Groep III wordt gevormd door de semantische differentiaal (zie 4.2.4.), en groep VII door schalen uit de ABV (zie 4.2.2.). Groep IV mikt op een 'zelfcritiek'-factor, groep V op een samenwerkingsfactor, groep VI op een prestatiefactor en groep VIII op een schoolattitudefactor (zie 4.2.5.). Hier en daar overlappen de groepen elkaar enigszins. Dit is gebeurd omdat de verwachting er was dat per groep, door de combinatie met andere variabelen, toch verschillende zinvolle factoren naar voren zouden komen. Groep VIII bestaat uit slechts één variabele omdat hij niet duidelijk kon worden ingedeeld bij een van de andere zeven groepen.

De acht groepen waren als volgt samengesteld:

Groep I	Groep II	Groep III
. gewicht	. plate-tapping	. leren
. lengte	. sit and reach	. doorzetten
. schouderbreedte	. 50 m. shuttle run	. samenwerken
. bekkenbreedte	. bent arm hang	. vroeg opstaan
. omvang bovenarm	. vertical jump	. hobby's
. omvang dijbeen	. handgrip	. transpireren
. omvang kuit	. 12 min. loop	. sport
. polsbreedte	. W 170	. elkaar helpen
. kniebreedte	. FEV %	. rekenen
. huidplooi biceps		. inspanning
. huidplooi triceps		
. huidplooi subscapularis		
. huidplooi suprailiaca		
. gecorr. armdiameter		
. % vet		

Groep IV

- SW
- S.W.
- testattitude (T)

Groep V

- samenwerken
- elkaar helpen
- extraversie (E)
- SAGS O
- SAGS G

Groep VI

- leren
- rekenen
- transpireren
- vroeg opstaan
- doorzetten
- inspanning
- PMT-K P
- PMT-K F-
- PMT-K F+

Groep VII

- neuroticisme (N)
- somato-neuroticisme (NS)
- extraversie (E)
- masculinity-femininity (MF)

Groep VIII

- schoolattitude

Niet in de groepen opgenomen variabelen

Een aantal variabelen werd niet opgenomen in de boven beschreven conceptuele groepen. Het betreft de volgende:

1. de gecombineerde F-schaal van de PMT-K, omdat wij meting van F+ en F- voldoende achtten;
2. SAGS-R, omdat deze score niets anders is dan de ratio tussen SAGS O en SAGS G, twee variabelen die wel zijn gemeten;
3. ABV-ja, 4. ABV-nee, 5. ABV-totaal, 6. ABV-??, omdat over de interpretatie van meetresultaten op deze schalen nog te weinig bekend is en ze bovendien niet van toepassing lijken te zijn op onze vraagstelling;
7. alle P- en A-factoren van de Standaard Semantische Differentiaal (SSD), omdat wij op goede gronden (zie 4.2.) mogen aannemen dat vooral de E-factor voor ons van belang is;
8. gymnastiek voorkeur en 9. gymnastiek rangnummer. Deze twee variabelen konden niet duidelijk in een van de acht groepen worden ingedeeld. Ze zijn opgenomen in een ongereduceerde multivariate analyse.

Ten aanzien van de gymnastiekprestaties zij vermeld dat de gymnologische variabelen niet in een der groepen is opgenomen omdat deze variabelen geheel vallen buiten de voortest-natest opzet en bovendien meerdere malen zijn gemeten gedurende de experimentele periode (zie 6.6.).

Statistische consequenties van verschillen in factorstructuur

Vooruitlopend op een nadere bespreking van de resultaten van de factoranalyses over groepen van variabelen (zie 7.2.), moet worden opgemerkt dat de correlatiestructuur van de variabelen in sommige van de acht groepen wezenlijk verschillend bleek te zijn omdat er geen voor experimentele en controlegroep gelijkgeldende factoren bleken te bestaan. Met name geldt dit voor de groepen III, V en VI. Besloten werd groep III in zijn totaliteit als een multivariaat tweesteekproevenprobleem ongereduceerd te analyseren omdat alle variabelen in deze groep tot hetzelfde instrument behoren. Ten slotte werd besloten de variabelen SAGS O, SAGS G, PMT-K P, PMT-K F-, PMT-K F+, Gymnastiek voorkeur en Gymnastiek rangnummer tezamen eveneens als een ongereduceerd multivariaat tweesteekproevenprobleem te analyseren (zie 6.4.).

6.3. COVARIANTIEANALYSE

Bij elke op grond van de factoranalyses verkregen representant werd gespecificeerd welke concomitanten naast de onafhankelijke variabele van invloed worden geacht, alsmede welke combinaties van concomitanten en onafhankelijke variabele konden bijdragen tot de verschilscores van deze representant.

De toename op een bepaalde variabele wordt bij voorbeeld niet veroorzaakt door de onafhankelijke variabele alleen, ook niet zozeer door buitenschoolse lichamelijke activiteiten alleen, maar veeleer door de combinatie van de onafhankelijke variabele en buitenschoolse activiteiten.

Het gaat er daarbij niet zozeer om de invloed van deze variabelen als zodanig te kwantificeren, als wel om hun uitwerking te scheiden van die van de onafhankelijke variabele en daardoor deze

laatste duidelijker zichtbaar te maken.

Aanvankelijk werd gedacht aan een 'matched comparison' techniek, dat wil zeggen, een vergelijking zou worden uitgevoerd tussen groepen proefpersonen die 'niet teveel verschillen' op concomitante variabelen. Omdat een redelijke operationalisatie van het begrip 'niet teveel verschillen' ondoenlijk leek, werd besloten de oplossing te zoeken door middel van covariantieanalyse (Scheffé, 1959). Op deze wijze kon rekening gehouden worden met de invloed van deze concomitanten.

Het is van belang het aantal concomitanten en interacties daarvan, die in het model gespecificeerd worden, zo klein mogelijk te houden om de nauwkeurigheid van de schattingen van de modelparameters niet te zeer te laten verminderen. Het volgende uitgangspunt is aangehouden: slechts wanneer een duidelijke theoretische argumentatie kan worden aangevoerd voor het verwachten van enig effect van een bepaalde concomitant, is deze in het model opgenomen (zie 7.3.). Dit had tot resultaat dat in alle gevallen buitenschoolse lichamelijke activiteit en de combinatie hiervan met de onafhankelijke variabele van invloed werden geacht, terwijl in een groot aantal gevallen bovendien de biologische leeftijd, maar niet de combinatie ervan met de onafhankelijke variabele een rol werd toegekend. Voor een derde concomitant, het lesgeefgedrag der leraren lichamelijke oefening (zie 5.3.), geldt dat deze variabele niet aan de proefpersonen zelf gemeten is, zodat hij niet in de covariantieanalyse kon worden opgenomen.

De volgende variabelen waren aan alle proefpersonen gemeten:

$z_j^{(e)}$: de verschilscore op een bepaalde variabele van proefpersoon j uit de experimentele groep;

$z_j^{(c)}$: de verschilscore op een bepaalde variabele van proefpersoon j uit de controlegroep;

$a_j^{(e)}$: de buitenschoolse lichamelijke activiteit, gemeten als totale stappenscore van proefpersoon j uit de experimentele groep;

$a_j^{(c)}$: idem, uit de controlegroep;

$l_j^{(e)}$: de biologische leeftijd gemeten op de voortest van proefpersoon j uit de experimentele groep;

$1_j^{(c)}$: idem, uit de controlegroep.

We beschouwen het volgende covariantieanalysemodel:

Model 1 :

$$Z_j^{(c)} = \mu + \alpha + \beta^{(c)}.a_j^{(c)} + \gamma.1_j^{(c)} + \Omega_j^{(c)} \qquad (j = 1,\ldots,37)$$

$$Z_j^{(e)} = \mu - \alpha + \beta^{(e)}.a_j^{(e)} + \gamma.1_j^{(e)} + \Omega_j^{(e)} \qquad (j = 1,\ldots,33)$$

Daarin is:

μ : het algemeen gemiddelde van deze variabele;

α : het effect van de onafhankelijke variabele:
+ α voor de controlegroep (3 lessen lichamelijke oefening)
- α voor de experimentele groep (5 lessen lichamelijke oefening)

$\beta^{(c)}$: het effect van één eenheid buitenschoolse lichamelijke activiteit méér in de controlegroep;

$\beta^{(e)}$: idem, in de experimentele groep;

γ : het effect van één eenheid biologische leeftijd méér, voor beide groepen gelijk;

$\Omega_j^{(c)}$: de storingsterm, dat wil zeggen, dat gedeelte van de verschilscore dat niet tot stand komt door de onafhankelijke variabele, buitenschoolse lichamelijke activiteit of biologische leeftijd, voor proefpersoon j in de controlegroep;

$\Omega_j^{(e)}$: idem, in de experimentele groep.

In dit model worden drie toetsen uitgevoerd:

Toets 1: nulhypothese: $\alpha = \beta^{(e)} = \beta^{(c)} = \gamma = 0$

alternatief: tenminste één van deze parameters is niet gelijk aan nul.

Wanneer hier de nulhypothese wordt verworpen, is er reden om aan te nemen dat tenminste één van de drie: onafhankelijke variabele, buitenschoolse lichamelijke activiteit,en/of biologische leeftijd, van

invloed is op de verschilscores op deze variabelen.

Toets 2: nulhypothese: $\alpha = 0$ én $\beta^{(e)} = \beta^{(c)}$

alternatief : $\alpha \neq 0$ óf $\beta^{(e)} \neq \beta^{(c)}$

Dit is in feite de toets waar het in het onderzoek om gaat. Wordt deze nulhypothese verworpen, dan concluderen wij dat de onafhankelijke variabele van invloed is, hetzij direct ($\alpha \neq 0$), hetzij in combinatie met buitenschoolse lichamelijke activiteit ($\beta^{(e)} \neq \beta^{(c)}$), hetzij beide.

Toets 3: nulhypothese: $\beta^{(e)} = \beta^{(c)}$

alternatief : $\beta^{(e)} \neq \beta^{(c)}$

Deze toets onderzoekt of er een apart combinatieeffect van de onafhankelijke variabele en buitenschoolse lichamelijke activiteit bestaat, dat wil zeggen, of de onafhankelijke variabele een andere uitwerking heeft op leerlingen met een hoog niveau aan buitenschoolse lichamelijke activiteit, dan op leerlingen met een laag niveau op deze concomitant.

De toetsingen vinden steeds plaats met F-toetsen (variantie-ratiotoetsen), onder de veronderstelling dat de Ω_j's onderling onafhankelijk en normaal verdeelde stochastische grootheden zijn, met verwachting nul, en voor elke proefpersoon gelijke variantie. Op te merken valt dat, wanneer de normaliteitsveronderstelling niet volkomen terecht is, de robuustheid van de F-toets ervoor zorg draagt dat de werkelijke onbetrouwbaarheid van de toetsingsprocedure niet veel afwijkt van de opgegeven onbetrouwbaarheid. Wel kan het onderscheidingsvermogen eronder lijden.

6.4. NIET GEREDUCEERD TWEESTEEKPROEVENPROBLEEM

Voor de groepjes waar de factoranalyse geen gelijkluidende sttructuur in experimentele en controlegroep opleverde, is verantwoorde reductie tot een beperkt aantal representerende variabelen niet mogelijk. In dergelijke gevallen is besloten tot ongereduceerd oplossen van het tweesteekproevenprobleem. Laten $(X_1,\ldots,X_n)$ de verschilscores van een persoon uit de experimentele groep zijn en $(Y_1,\ldots,Y_n)$

die van een persoon uit de controlegroep (n = het aantal variabelen in een groepje). De beantwoording van de vraag: 'Heeft de onafhankelijke variabele effect op deze variabelen?' wordt nu gezocht in het toetsen van de nulhypothese $H_0 : E(X_1,...,X_n) = E(Y_1,...,Y_n)$ tegen het alternatief dat deze verwachtingen ongelijk zijn. Mits wij aannemen dat $(X_1,...X_n)$, respectievelijk $(Y_1,...,Y_n)$ elk een multivariaatnormale verdeling hebben met gelijke covariantiematrices, kunnen we deze hypothese toetsen met behulp van Hotelling's tweesteekproeven T^2-toets (Morrison, 1967). Deze toets is altijd tweezijdig en concomitanten kunnen er niet in worden opgenomen. De robuustheid van deze toets garandeert dat de toetsing valide is, ook in geval van afwijking van multivariate normaliteit. De gelijkheid van de covariantiematrices is een cruciale veronderstelling die hier wel enige vraagtekens oproept. De toets wordt in dit onderzoek immers toegepast wanneer op de voortest geen gelijke correlatiestructuur wordt aangetroffen. Hoewel dit niet a priori impliceert dat ook de covariantiestructuur van de verschilscores ongelijk is, is enige argwaan wel op zijn plaats. Blijkt de toets H_0 te verwerpen, dan geldt dat met een onbetrouwbaarheid van 5% is aangetoond dat voor minstens één variabele, $E X_i$ en $E Y_i$ ongelijk zijn. In zulk een geval wil men graag weten welke variabele(n) dan wel zo een verschil vertoont (nen). Dit kan onderzocht worden door betrouwbaarheidsintervallen voor elk van de verschillen $E X_i - E Y_i$,i = 1,...,n op te stellen, wier gezamenlijke onbetrouwbaarheden een voorgeschreven onbetrouwbaarheid van 5% niet overschrijven, en dan te zien welke van deze intervallen het getal nul niet bevat. Het is evenwel niet zo dat wij dan altijd een of meer variabelen kunnen vinden waar een verschil kan worden aangetoond. Soms verschillen alle variabelen elk slechts een niet aantoonbaar klein beetje, terwijl het totale verschil voor alle tezamen wel significant blijkt.

6.5. VERGELIJKING VAN DE UITGEVOERDE ANALYSE MET EEN ANDERE METHODE

In veel sociaalwetenschappelijk onderzoek wordt een voortest-natest opzet geanalyseerd door voor alle betrokken variabelen een tweesteekproeven t-toets uit te voeren. Een dergelijke methode negeert

echter het essentieel multivariate karakter van het probleem volledig. Bij dergelijke t-toetsen wordt een onbetrouwbaarheidsdrempel van 5% aangehouden, dat wil zeggen dat in lange reeksen onafhankelijke toepassingen van de toets, wanneer de nulhypothese waar is, met een frequentie van niet meer dan een op de twintig het foutieve resultaat: verwerp de nulhypothese, gegeven wordt. Hoe de kans op ten onrechte verwerpen van nulhypothesen zich gedraagt bij afhankelijke toepassingen - het gaat immers om variabelen die aan dezelfde proefpersonen zijn gemeten, en dus afhankelijk kunnen zijn - is in zijn algemeenheid nogal onduidelijk, evenwel is het zeker niet denkbeeldig dat de kans dat men een tweede nulhypothese ten onrechte verwerpt wanneer men reeds een nulhypothese abusievelijk heeft verworpen, veel groter uitvalt dan de voorgeschreven waarde van 5%. De fractie foutieve uitspraken onder alle uitspraken betrekking hebbende op vergelijkingen waar de nulhypothese (geen verschil) in feite waar is, kan dan veel groter worden dan de getolereerde eentwintigste. Dergelijke cumulaties van fouten tellen des te zwaarder naarmate er meer vergelijkingen worden uitgevoerd. In de in dit onderzoek toegepaste methode wordt echter expliciet rekening gehouden met de afhankelijkheid van de betrokken variabele. Ook nu nog zijn meerdere, mogelijk afhankelijke, toetsen uitgevoerd, maar hun aantal blijft beperkt, terwijl daarenboven moeite is gedaan om toetsen uit te voeren op althans conceptueel min of meer onafhankelijke variabelen.

Het bovenstaande vormt niet alleen de rationale achter de multivariate aanpak, het is tevens een critische kanttekening bij tot op heden veelvuldig gehanteerde procedures bij een groot deel van het sociaalwetenschappelijk onderzoek, waarbij op een zeer groot aantal variabelen en weinig proefpersonen lange reeksen afhankelijke toepassingen van de t-toets worden gebezigd en groepen als gelijk (ongelijk) worden gekwalificeerd op basis van gelijke (ongelijke) gemiddelden alléén op een aantal variabelen, zonder dat men enige notie heeft van overeenkomst in correlatiestructuur tussen de variabelen.

6.6. ANALYSE VAN GYMNOLOGISCHE VARIABELEN

Buiten de voortest-natest opzet om werden nog een aantal extra

gegevens verzameld. Het gaat hier om de gymnologische variabelen die gedurende de experimentele periode op vier tijdstippen werden gemeten (zie 4.1.3.). Op ieder tijdstip waren de tests verschillend. Om na te gaan of de gymnastiekprestaties gedurende het jaar meer toegenomen zijn bij de experimentele groep dan bij de controlegroep, is de volgende analyse uitgevoerd.

Laten (X_1 , X_2 , X_3 , X_4) de scores zijn op tests gedurende de vier meettijdstippen voor de experimentele groep en (Y_1 , Y_2 , Y_3 , Y_4) idem voor de controlegroep. $E X_i - E Y_i = v_i$, zijn de verschillen in verwachting over de verdeling van scores tussen beide proefgroepen op tijdstip i. Getoetst wordt eerst de nulhypothese: $H_0 : E X_i = E Y_i$, of anders geformuleerd: $v_i = 0$, voor i = 1,2,3,4, tegen het alternatief H_1: minstens één $v_i \neq 0$. Hier wordt dus getoetst of er wel verschil is tussen beide proefgroepen. De gebruikte toets is Hotelling's T^2 (Morrison, 1967).

Bovendien is nagegaan of de verschillen met de tijd steeds groter worden. Wanneer er sprake is van zich niet wijzigende of onregelmatig verlopende verschillen, is de expressie $(-9v_1 + 2v_2 + 3v_3 + 4v_4)$ ongeveer nul, terwijl als er regelmatig toenemende verschillen bestaan, deze expressie positief is. Om een lineaire trend aan te tonen is daarom de nulhypothese: H_0: $(-9v_1 + 2v_2 + 3v_3 + 4v_4) = 0$, tegen het alternatief: H_1: $(-9v_1 + 2v_2 + 3v_3 + 4v_4) > 0$ getoetst. De trendanalyse is uitgevoerd door bij Hotelling's T^2 een onbetrouwbaarheidsinterval voor $(-9v_1 + 2v_2 + 3v_3 + 4v_4)$ op te stellen.

Literatuur

1. Barrow, H.M., R. Mc Gee, A practical approach to measurement in physical education, Lea en Febiger, Philadelphia, 1968.
2. Beunen, G., Studie in het verband tussen twee onafhankelijke estimaties van de botleeftijd bij jongens, Hermes, 4, 1970, 117-182.
3. Conover, W.J. , Practical nonparametric statistics, Wiley, New-York, 1971.
4. Flanders, N.A., Interaction Analysis in the Classroom: A Manual for Observers, (mimeographed), 1960.
5. Flanders, N.A., E.J. Amidon, J.B. Hough, Interaction Analysis: Theory, Research and Application, Addison-Wesley, Massachusetts, 1967, 158.
6. Groot, A.D. de, Methodologie, Mouton en Co. Den Haag, 1968.
7. Jonge, H. de, Inleiding tot de medische statistiek, N.I.P.G., Leiden, 1963.
8. Jöreskog, K.G., Statistical Estimation in Factor Analysis, Almqvist en Wiksell, Stockholm, 1963.
9. Kemper, H.C.G., J.J.L. Pieters, Comparative study of anthropometric measurements of the same subjects in two different institutes, Am. J. of Phys. Anthrop., 1973 (in press).
10. Kendall, M.G., Rank Correlation Methods, Griffin, London, 1962.
11. Medicus, H., A. Grøn, C.F.A. Moorrees, Reproducibility of rating stages of osseous development (Tanner-Whitehouse system), Am. J. Phys. Anthrop., 35, 1971, 359-372.
12. Morrison, D.F., Multivariate Statistical Methods, McGraw-Hill, New-York, 1967.
13. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor VWO en HAVO op lichamelijke en geestelijke groei en ontwikkeling. Doctoraal werkstuk, Lab. voor Psychofysiologie, Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv., 59, 12 en 13, 1971, 543 en 604.
14. Scheffé, H., The Analysis of Variance, Wiley, New York, 1959.
15. Scott, W.A., The case of nominal scale coding, Public Opinion Quarterly, 1955, 321-325.

16. Siegel, S., Nonparametric Statistics for the behavioral sciences, McGraw-Hill, New York, 1956.

hoofdstuk 7

HYPOTHESE VORMING

7.1. VOORLOPIGE HYPOTHESEN

- hypothesen met betrekking tot de antropometrische variabelen
- hypothesen met betrekking tot de fysiologische variabelen
- hypothesen met betrekking tot de gymnologische variabelen
- hypothesen met betrekking tot de psychologische variabelen

7.2. RESULTATEN FACTORANALYSE

- groep I, de antropometrische variabelen
- groep II, de fysiologische variabelen

- groep III, concepten beoordeeld met de standaard semantische differentiaal
- groep IV, sociale wenselijkheid en testattitude
- groep V, de concepten samenwerken en elkaar helpen, de extraversiescore en de S.A.G.S.
- groep VI, variabelen die verband houden met prestaties, doorzettingsvermogen en sport
- groep VII, variabelen uit de A.B.V.
- groep VIII, schoolattitude

7.3. HYPOTHESEN GETOETST DOOR MIDDEL VAN COVARIANTIE ANALYSE

- storende variabelen
- interacties

7.4. HYPOTHESEN GETOETST DOOR MIDDEL VAN EEN NIET GEREDUCEERDE MULTIVARIATE ANALYSE

7.5. HYPOTHESE MET BETREKKING TOT DE GYMNOLOGISCHE VARIABELEN

LITERATUUR

7. HYPOTHESE VORMING

7.1. VOORLOPIGE HYPOTHESEN

Op grond van de in het onderzoek gehanteerde afhankelijke variabelen (zie 4.) zijn verwachtingen geformuleerd over de invloed daarop van de onafhankelijke variabele, die door middel van de onderstaande voorlopige hypothesen kunnen worden getoetst.

Hypothesen met betrekking tot de antropometrische variabelen

- de skeletgrootte factor (lengte, schouderbreedte) neemt toe bij de experimentele groep ten opzichte van de controlegroep;
- de botbreedte factor (pols- en kniebreedte) neemt toe bij de experimentele groep ten opzichte van de controlegroep;
- de vetfactor (% vet voorspeld uit twee huidplooien) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de spierfactor (omvang kuit en gecorrigeerde armdiameter) neemt bij de experimentele groep toe ten opzichte van de controlegroep.

Hypothesen met betrekking tot de fysiologische variabelen

- de duurkracht (bent arm hang) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de explosieve armkracht (handgrip) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de explosieve beenkracht (vertical jump) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de lenigheid (sit and reach) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de snelheid van voortbewegen (50 meter shuttle run) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de snelheid van ledematen (plate tapping) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- het algemene uithoudingsvermogen (12 minuten loop) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de W_{170} (het arbeidsvermogen dat bij een hartfrequentie van 170 slagen/minuut geleverd kan worden per kilogram lichaamsgewicht) neemt bij de experimentele groep toe ten opzichte van de

controlegroep.

- de prestatie van de ventilatie (FEV %) neemt bij de experimentele groep toe ten opzichte van de controlegroep.

Hypothese met betrekking tot de gymnologische variabelen

- de gymnastiekscore (samengesteld uit een aantal gymnologische tests) neemt bij de experimentele groep meer toe dan bij de controlegroep.

Hypothesen met betrekking tot de psychologische variabelen

- de N score (ABV) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de NS score (ABV) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de E score (ABV) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de T score (ABV) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de P score (PMT K) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de F- score (PMT K) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de F+ score (PMT K) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de S.W. score (PMT K) neemt bij de experimentele groep af ten opzichte van de controlegroep;
- de factor samenwerken (SSD) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de factor leerprestatie (SSD) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de factor recreatie en vrijetijdsbesteding (SSD) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de factor prestatiebereidheid (SSD) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de factor prestatiegeschiktheid (SSD) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- het gemiddelde van de gegeven scores (SAGS) neemt bij de experimentele

groep toe ten opzichte van de controlegroep;
- de attitude ten aanzien van de school (schoolattitudetest) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
- de voorkeur ten aanzien van het vak gymnastiek (schoolvoorkeurs-enquête, rangnummer gymnastiek) neemt bij de experimentele groep toe ten opzichte van de controlegroep.

7.2. RESULTATEN FACTORANALYSE

De factoranalyse, toegepast op gegroepeerd materiaal om daarmee reductie van het aantal variabelen te verkrijgen (zie 6.2.), laat de volgende resultaten zien:

Groep I, de antropometrische variabelen

Ten aanzien van deze variabelen werden in beide groepen duidelijk drie factoren gevonden: 1. een 'vet' factor; 2. een 'skelet en grootte' factor; 3. een 'spier' factor (zie tabel 7.1.).

Deze factoren werden eveneens gevonden in het vooronderzoek bij een vergelijkbare populatie proefpersonen (Oud e.a., 1971). Naast een 'skelet en groottte' factor werd toen echter ook een 'botbreedte' factor afgezonderd.

De variabelen met de hoogste factorladingen, genomen over de totale groep, waren:
ad 1. het percentage vet;
ad 2. gewicht (lengte);
ad 3. gecorrigeerde armdiameter.

Hoewel een 'skelet en grootte' factor kon worden aangetoond waarin lengte en gewicht als variabelen naar voren kwamen met de hoogste factorladingen, zal deze factor bij de hypothesevorming desondanks buiten beschouwing blijven, omdat deze niet voldoende onafhankelijk is van de andere factoren. Zowel lengte als gewicht vertonen zeer hoge correlaties met variabelen van de 'vet en spier' factor.

Groep II, de fysiologische variabelen

Ten aanzien van de lichamelijke prestatietests werden de volgende factoren gevonden (tussen haakjes staan vermeld de variabelen met

variables \ factors	total groups					
	1	2	3	4	5	6
weight		+				
height		+				
biacromial diameter		+				
biiliocristal diameter		+				
upper arm circumference	+					
thigh circumference	+					
calf circumference		+				
wrist breadth		+				
bycondylar femur		+				
bicipital skinfold	+					
tricipital skinfold	+					
subscapular skinfold	+					
suprailiacal skinfold	+					
corr. upper arm diameter				+		
% fat	+					

Tabel 7.1. : Factor loadings of the anthropometric variables in a 6-factor model, represented from the total groups (n = 70);
(+ or - means a factor loading higher than + or - .60).

de hoogste factorladingen):

1. snelheid van de ledematen (plate tapping);
2. snelheid van voortbewegen (50 meter shuttle run);
3. duurkracht (bent arm hang);
4. explosieve armkracht (handgrip);

verder kramen als aparte factoren naar voren:

5. algemeen uithoudingsvermogen (W_{170});
6. ventilatieprestatie (FEV %), (zie tabel 7.2.).

De factor explosieve armkracht (handgrip) komt niet voor in de totale groep. Toch is deze gekozen als aparte factor, gezien het feit dat deze factor in de experimentele en controlegroep afzonderlijk wel voorkomt.

factors / variables	control group							experimental group							total group						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
plate tapping			+						-								-				
sit and reach										+						-					
50 meter shuttle run	+														+						
bent arm hang					+			-													+
vertical jump						+														+	
handgrip							-					+									
12 min. run walk	-							-													
W_{170}			-											-					-		
F.E.V. %		-									+							-			

Tabel 7.2. : Factor loadings of the physiological variables in a 7-factor model, represented from the control, experimental and total group (n = 37,33 and 70); (+ or - means a factor loading higher than + or - .60).

De 12-minutenloop verviel als maat voor het algemeen uithoudingsvermogen omdat deze nergens een aparte factor vormde en omdat de correlatie met de W_{170}, waarvan bekend is dat het een betrouwbare maat is voor het bepalen van het algemeen uithoudingsvermogen, zelfs lager is dan met andere tests, waarvan vaststaat dat ze geen maat zijn voor het algemeen uithoudingsvermogen.

Ook de 'vertical jump' en de 'sit and reach' vervielen. Van deze tests is bekend dat ze bij 12 en 13-jarigen niet altijd even sterk aanwezig zijn (zie 4.1. 2.) In dit onderzoek was bij de voortest de 'sit and reach ' wel, en de 'vertical jump' niet een aparte factor bij de experimentele groep. Het omgekeerde was het geval bij de controlegroep. Bovendien bleek dat bij zes van de zeven verdachte correlatieverschillen tussen de experimentele en de controlegroep (absoluut verschil groter dan 0,40) deze tests betrokken waren.

Groep III, concepten beoordeeld met de standaard semantische differentiaal

De factorstructuur ten aanzien van de tien concepten beoordeeld met de standaard semantische differentiaal wat betreft de evaluatieve dimensie, leverde zulke grote verschillen op tussen de experimentele en controlegroep, en de samenhang tussen de concepten was dermate zwak, met bovendien een geringe bijdrage aan de variantie (49,27%), dat niet een aantal voor beide groepen gelijke factoren kon worden afgezonderd (zie tabel 7.3.).

factors / variables	control group					experi-mental group					total group				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
learning															
perseverance	-										-				
cooperating		-					-								
getting up early					+										+
hobbies		-				-						-			
perspiring			-												
sports	-														
helping each other						-									
arithmatics				-											
exertion								+							

Tabel 7.3. : Factor loadings of the 10 concepts, judged with the standard semantic differential, in a 5-factor model, represented from the control, experimental and total group (n = 37, 33 and 70);
(+ or - means a factor loading higher than + or - .60).

De verschillen tussen de experimentele en controlegroep betroffen vooral de concepten hobby's, vroeg opstaan, transpireren en elkaar helpen. Verder bleek dat de factorstructuur, gevonden in de gegevens van het vooronderzoek (intern rapport), in dit experiment totaal niet terug te vinden was (zie 4.2.).

Uit de correlatieverschillen tussen beide groepen bleek dat alle verschillen van boven de 0,40 wegvielen als de concepten hobby's en

zouden worden geëlimineerd.

Groep IV, sociale wenselijkheid en testattitude

Bij de toepassing van een éénfactor model bestond er tussen de experimentele en controlegroep in zoverre overeenkomst, dat alle variabelen redelijke ladingen vertoonden op de factor (minstens 0,40 en alle negatief). Wel vertoonde de correlatie tussen de SW van de PMT K en de testattitude van de ABV een duidelijk verschil tussen de experimentele en controlegroep (0,30).

Dit heeft geleid tot het besluit de testattitude van de ABV te nemen, niet als representant van deze groep, maar als zelfstandige variabele uit genoemde test en de twee sociale wenselijkheidsschalen alleen te gebruiken bij de interpretatie van de ABV en de PMT K.

Groep V, de concepten samenwerken en elkaar helpen, de extraversiescore en de SAGS

De factorladingen van de concepten samenwerken en elkaar helpen waren tegengesteld van teken wat betreft de experimentele en controlegroep, en ook de correlatiematrices van beide groepen vertoonden nogal grote verschillen (zie tabel 7.4.).

In verband met het aantal uitvallers op grond van criteria die te maken hebben met de onafhankelijke variabele, is het de vraag of de SAGS daardoor niet minder bruikbaar is geworden omdat de klas als geheel niet intact blijft (zie 2.1. en 9.6.)

Het bovenstaande had evenwel tot gevolg dat niet tot een overeenkomstige factorstructuur voor de experimentele en controlegroep kon worden besloten.

Groep VI, variabelen die verband houden met prestaties, doorzettingsvermogen en sport

In deze groep valt geen duidelijke overeenkomst in factorstructuur waar te nemen tussen de experimentele en controlegroep. Vooral het verschil in hoogte van de ladingen van dezelfde variabele is vaak aanzienlijk. Hierdoor bleek het onmogelijk om voor de experimentele

factors / variables	control group		experi-mental group		total group	
	1	2	1	2	1	2
cooperating			+			
helping each other						
extraversion						
scores received SAGS(O)	-				+	
scores given SAGS (G)	-				+	

Tabel 7.4. : Factor loadings of the concepts cooperating and helping each other, the extraversion score and the variables received and given judgements of the SAGS in a 2-factor model represented from the control, experimental and total group (n = 37, 33 and 70);
(+ or - means a factor loading higher than + or - .60).

en controlegroep tot de keuze van eenzelfde representatieve variabele te komen (zie tabel 7.5.).

factors / variables	control group			experi-mental group			total group		
	1	2	3	1	2	3	1	2	3
learning	-					-		-	
arithmatics									
perspiring									
getting up early				-					
perseverance	-			-					
exertion				-		-	+		
PMT K P	-				-			-	
PMT K F-		-			+				
PMT K F+									

Tabel 7.5. : Factor loading of six concepts related to performances and/or perseverance and the PMT K P, F- and F+ score in a 3-factor model, represented from the control, experimental and total group (n = 37, 33 and 70);
(+ or - means a factor loading higher than + or - .60).

Groep VII, variabelen uit de ABV (N, NS, E en MF)

De factorstructuur is voor de experimentele en controlegroep vrijwel gelijk (zie tabel 7.6.). De variabele met de hoogste lading voor beide groepen is de N score (neuroticisme). Deze bedraagt 0,78

factors / variables	control group		experi-mental group		total group	
	1	2	1	2	1	2
neuroticism	+		+		+	
somato neuroticism	+		+		+	
extraversion		-				-
masculine/feminine score		-				-

Tabel 7.6. : Factor loadings of four variables from the ABV in a 2-factor model represented from the control, experimental and total group (n = 37, 33 and 70); (+ or - means a factor loading higher than + or - .60).

over de totale groep. Voor de vraagstelling van het onderzoek zou de NS score te prefereren zijn gezien de relatie met psychosomatische klachten, maar op grond van de geringe betrouwbaarheid van deze score in de leeftijd van dertien tot negentien jaar (0,39; zie Wilde, 1970), is besloten de N score te kiezen. Tevens zal de extraversie-score opgenomen worden als representant voor zijn gebied.

Groep VIII, de schoolattitude

Omdat in deze groep slechts één variabele was opgenomen, werd geen factoranalyse uitgevoerd. Bij vergelijking van deze variabele met de andere psychologische variabelen blijken er geen al te grote correlatieverschillen waar te nemen tussen experimentele en controlegroep. Op grond hiervan achten wij het geoorloofd deze variabele op te nemen in de covariantieanalyse samen met de hierboven gevonden variabelen, die zowel voor de experimentele als controlegroep gekozen konden worden, gezien hun overeenkomst in factorstructuur.

7.3. HYPOTHESEN, GETOETST DOOR MIDDEL VAN COVARIANTIEANALYSE

Op grond van de resultaten uit de factoranalyse konden de voorlopige hypothesen tot onderstaande hypothesen gereduceerd worden:

. met betrekking tot de antropometrische variabelen:

1. de vet factor (% vet) neemt bij de experimentele groep af ten opzichte van de controlegroep;
2. de spier factor (gecorrigeerde armdiameter) neemt bij de experimentele groep toe ten opzichte van de controlegroep;

. met betrekking tot de fysiologische variabelen:

3. de duurkracht (bent arm hang) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
4. de explosieve armkracht (handgrip) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
5. de snelheid van het voortbewegen (50 meter shuttle run) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
6. de snelheid van de ledematen (plate tapping) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
7. het algemeen uithoudingsvermogen (W_{170}) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
8. de ventilatieprestatie (FEV %) neemt bij de experimentele groep toe ten opzichte van de controlegroep;

. met betrekking tot de psychologische variabelen:

9. de testattitude (T score ABV) neemt bij de experimentele groep af ten opzichte van de controlegroep;
10. de mate van neuroticisme (N score ABV) neemt bij de experimentele groep af ten opzichte van de controlegroep;
11. de mate van extraversie (E score ABV) neemt bij de experimentele groep toe ten opzichte van de controlegroep;
12. de schoolattitude (schoolattitudetest) neemt meer toe bij de experimentele groep ten opzichte van de controlegroep.

storende variabelen

Ten aanzien van de storende variabelen is vermeld dat deze in het covariantiemodel zouden worden opgenomen wanneer de verwachting bestaat dat die variabelen op enigerlei wijze mede het effect zouden

kunnen bepalen dat aan de onafhankelijke variabele zou worden toegeschreven (zie 6.3.).

In dit onderzoek zijn drie storende variabelen opgenomen (zie 5.):

1. de biologische leeftijd van de proefpersonen;
2. de buitenschoolse lichamelijke activiteit van de proefpersonen;
3. het lesgeefgedrag van de leraren.

De derde storende variabele werd niet in het model opgenomen omdat:

a. het lesgeefgedrag niet aan de proefpersonen zelf werd gemeten;
b. zowel de twee klassen uit de experimentele groep als uit de controlegroep van verschillende leraren les kregen, waardoor het leraar- en/of klaseffect niet te scheiden waren.

De argumentatie om de buitenschoolse lichamelijke activiteit als storende variabele op te nemen is de volgende: bij proefpersonen die in hun vrije tijd een hoge mate van lichamelijke activiteit te zien geven, mag een betere lichamelijke conditie verondersteld worden dan bij minder actieve proefpersonen. Het effect van slechts twee lessen extra lichamelijke oefening per week zal voor de eerstgenoemden veel geringer zijn, omdat bij beter 'getrainden' de trainingseffectiviteit afneemt zodat een intensiever trainingsprogramma moet worden opgezet om hetzelfde effect te hebben. Dit in tegenstelling tot de proefpersonen met een geringe mate van lichamelijke activiteit. Voor hen heeft een verhoging met twee lessen lichamelijke oefening per week relatief gezien veel meer betekenis; er mag dan ook een groter trainingseffect worden verwacht.

Ten aanzien van de biologische leeftijd is de argumentatie als volgt: op de leeftijd waarin deze proefpersonen zich bevinden, bestaan grote verschillen in biologische leeftijd (zie 5.1.). Op grond van het feit dat een grotere trainbaarheid zou mogen worden verwacht door de hogere concentratie van het groeihormoon in het bloed en van de androgenen tijdens de pubertaire groeispurt lijkt het aannemelijk dat het niveau van rijping aan het begin van de experimentele periode en verschillen in rijpingssnelheid tijdens de experimentele periode van invloed zullen zijn op de effecten van de onafhankelijke variabele.

Bovenstaande argumentatie is voor het onderzoek aanleiding

geweest tot het opnemen van de buitenschoolse lichamelijke activiteit als storende variabele bij alle twaalf hypothesen; de biologische leeftijd werd als storende variabele niet opgenomen bij de hypothese omtrent de testattitude, extraversie en schoolattitude, maar wel bij de overige negen hypothesen. Ten aanzien van de psychologische variabelen is de biologische leeftijd dus alleen bij de hypothese over de neuroticisme score opgenomen. De veronderstelling daarbij is dat de snelle ontwikkeling tijdens deze leeftijdfase vooral van invloed zou kunnen zijn op de geestelijke labiliteit van de proefpersonen, waardoor ten aanzien van de neuroticisme score een vertekend beeld zou kunnen ontstaan.

Interacties

Bij alle hypothesen zal tevens worden nagegaan of er effecten zijn die afhankelijke blijken te zijn van de interactie tussen de onafhankelijke variabele en de buitenschoolse lichamelijke activiteit. Dit is niet nagegaan ten aanzien van de interactie tussen de onafhankelijke variabele en de biologische leeftijd omdat het niveau van rijping aan het begin van de experimentele periode geen adequate informatie geeft om tot een zinvolle uitspraak over eventuele interactie-effecten te kunnen komen. Daarvoor is tevens nodig het niveau van rijping aan het einde van de experimentele periode en ook de mate van rijpingssnelheid gedurende deze periode.

Een keuze van een van deze drie mogelijkheden blijft het probleem geven van informatieverlening die niet volledig is te noemen. Een toevoeging van alle drie interactietoetsingen zou echter een te zware belasting voor het covariantieanalysemodel beteken (zie 6.3.); allemaal redenen om voorlopig af te zien van de interactietoetsing van de onafhankelijke variabele en de biologische leeftijd (zie schema 7.7.).

7.4. HYPOTHESEN, GETOETST DOOR MIDDEL VAN EEN NIET GEREDUCEERDE MULTIVARIATE ANALYSE

De factoranalyse die is uitgevoerd over de acht conceptuele groepen heeft niet voor iedere groep geleid tot reductie van het aantal

hypotheses	interfering variables		testing of interaction between habitual physical activity and the independent variable
	habitual physical activity	skeletal age	
fat mass	x	x	x
muscle mass	x	x	x
muscular endurance	x	x	x
explosive arm strength	x	x	x
running speed	x	x	x
speed of limb movement	x	x	x
muscular endurance	x	x	x
ventilation	x	x	x
test attitude	x		x
neuroticism	x	x	x
extraversion	x		x
school attitude	x		x

Schema 7.7. : Survey of the hypotheses with their interfering variables, and of testing interaction between habitual physical activity and the independent variable.

variabelen; de factorstructuur bleek namelijk in sommige gevallen in de experimentele en controlegroep niet overeen te stemmen (zie 7.2.). Het gevolg hiervan is dat een aantal variabelen toch door middel van een niet gereduceerde multivariate analyse getoetst moeten worden. Het gaat hier om de volgende variabelen:

1. de tien concepten, beoordeeld met de standaard semantische differentiaal;
2. de sociale wenselijkheidsscore van de ABV en de PMT-K;
3. de P, F- en F+ score van de PMT-K;
4. het gemiddelde van de ontvangen en gegeven scores van de SAGS.

De tien concepten die beoordeeld zijn met behulp van de standaard semantische differentiaal, zullen afzonderlijk van de overige variabelen op verschillen worden getoetst.

Het betreft de volgende concepten (groep A):

1. leren
2. doorzettingsvermogen
3. samenwerken
4. vroeg opstaan
5. hobby's
6. transpireren
7. sport
8. elkaar helpen
9. rekenen
10. inspanning

De reden voor deze afzonderlijke toetsing van de tien concepten is gelegen in het feit dat deze concepten in het vooronderzoek bij een vergelijkbare populatie proefpersonen de hoogst ladende variabelen op vijf factoren bleken te zijn (zie 4.2.) en dat in dit onderzoek deze factoren niet, of zeer zwak, te voorschijn kwamen (zie 7.2.). In het vooronderzoek bleek dat met name ten aanzien van de factor recreatie en vrijetijdsbesteding (met de concepten sport, hobby's en gymnastiek) de experimentele groep een significant hogere waardering uitsprak dan de controlegroep (Oud e.a., 1971).

Groep B bestond uit de volgende variabelen:

1. ontvangen scores (SAGS)
2. gegeven scores (SAGS)
3. prestatiemotivatie (PMT K)
4. negatieve faalangst (PMT K)
5. positieve faalangst
6. gymnastiek voorkeur
7. gymnastiek rangnummer

De variabelen gymnastiek voorkeur en gymnastiek rangnummer waren niet in de factoranalyse opgenomen (zie 6.2.); de variabelen zijn nu toegevoegd omdat ze een indruk geven over de plaats die het vak lichamelijke oefening inneemt tussen de andere schoolvakken.

De sociale wenselijkheidsscores werden niet opgenomen in het model omdat deze beide scores in hun respectievelijke tests dienst doen als controle op het wel of niet volgens eigen opvattingen invullen van de items uit de vragenlijsten.

In de niet gereduceerde multivariate analyse wordt gekeken naar de verschilscores van de voor- en natest tussen de experimentele en controlegroep. Er kunnen geen storende variabelen worden

opgenomen en de toets is altijd tweezijdig (zie 6.4.).

De hypothese voor zowel groep A als B luidt dan:

- de verschilscore is voor de experimentele groep en de controlegroep niet gelijk.

7.5. HYPOTHESE MET BETREKKING TOT DE GYMNOLOGISCHE VARIABELEN

De hypothese omtrent de gymnologische variabelen die door middel van tests meerdere malen gedurende de experimentele periode is afgenomen, luidt:

- de gymnastiekprestaties (gymnologische tests) zijn op de tijdstippen van meting bij de experimentele groep hoger dan bij de controlegroep.

Literatuur

1. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke oefening op leerlingen uit de brugklassen van een school voor VWO en HAVO op lichamelijke en geestelijke groei en ontwikkeling. Werkstuk Laboratorium voor Psychofysiologie, Universiteit van Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv., 59, 12 en 13, 1971, 543 en 604.
2. Wilde, G.J.S., Neurotische labiliteit, gemeten volgens de vragenlijstmethode, T. v. Rossen, Amsterdam, 1970.

hoofdstuk 8

RESULTATEN

8.1. RESULTATEN COVARIANTIEANALYSE

8.2. RESULTATEN NIET GEREDUCEERDE MULTIVARIATE ANALYSE

8.3. RESULTATEN GYMNOLOGISCHE VARIABELEN

8.1. RESULTATEN COVARIANTIEANALYSE

Bij de toetsing van de in hoofdstuk 7.3. opgestelde hypothesen door middel van covariantieanalyse (zie 6.3.) is getracht een effect aan te tonen van de onafhankelijke variabelen, waarbij de buitenschoolse lichamelijke activiteit en/of de biologische leeftijd werden beschouwd als storende variabelen. Ten aanzien van de buitenschoolse lichamelijke activiteit geldt verder nog dat eveneens werd nagegaan of deze storende variabele in combinatie met de onafhankelijke variabele van invloed is geweest. De biologische leeftijd werd niet in combinatie met de onafhankelijke variabele op zijn invloed onderzocht.

Zoals onder 6.3. uitgebreid werd beschreven, zijn per hypothese een aantal toetsen uitgevoerd:

Toets 1 : de nulhypothese kon slechts in twee van de twaalf gevallen worden verworpen, namelijk bij de factor spier en bij de factor explosieve armkracht. Dit betekent dat alleen bij deze twee variabelen reden is om aan te nemen dat er een effect bestaat, veroorzaakt door de onafhankelijke variabele ($\alpha \neq 0$) en/of buitenschoolse lichamelijke activiteit ($\beta^{(e)} \neq \beta^{(c)}$) en/of biologische leeftijd ($\gamma \neq 0$).

Alleen bij de factor spier en bij de factor explosieve armkracht heeft het dus zin de toetsen 2 en 3 uit te voeren.

Toets 2 : de nulhypothese kon alleen worden verworpen ten aanzien van explosieve armkracht, hetgeen impliceert dat er reden is om aan te nemen dat een effect bestaat van de onafhankelijke variabele, óf alleen, óf in combinatie met buitenschoolse lichamelijke activiteit (respectievelijk $\alpha \neq 0$ en $\beta^{(e)} \neq \beta^{(c)}$). In toets 3 zal worden nagegaan of de buitenschoolse lichamelijke activiteit het effect op de explosieve armkracht mede heeft veroorzaakt.

Toets 3 : de nulhypothese ($\beta^{(e)} = \beta^{(c)}$) kon niet worden verworpen, hetgeen betekent dat de onafhankelijke variabele alleen van invloed is op de explosieve armkracht en wel op 1%-niveau.

Ten aanzien van de factor spier kan het volgende worden opgemerkt: Toets 1 gaf aan dat het zinvol was verder te gaan zoeken naar mogelijke effecten van de onafhankelijke variabele en/of buitenschoolse

factor	variable	test 1 $H_0: \alpha = \beta^{(e)} = \beta^{(c)} = \gamma = 0$	test 2 $H_0: \alpha = 0$ and $\beta^{(e)} = \beta^{(c)}$	test 3 $H_0: \beta^{(e)} = \beta^{(c)}$	remarks
ventilation	FEV %	n.s.			-
muscle mass	corr. upper arm diam.	*	n.s.		
fat mass	% fat	n.s.			0
speed of limb movement	plate tapping	n.s.			
running speed	50 m. shuttle run	n.s.			0
muscular endurance	bent arm hang	n.s.			-
explosive arm strength	handgrip	**	**	n.s.	
aerobic power	W_{170}	n.s.			0
school attitude	stand.score $(-\gamma)$	n.s.			
neuroticism	N score	n.s.			0
extraversion	E score $(-\gamma)$	n.s.			-
selfdefensiveness	T score $(-\gamma)$	n.s.			0

Tabel 8.1. : Results of the analysis of covariance: H_0 = null hypothesis, α = the independent variable, β = habitual physical activity, γ = skeletal age; n.s. = not significant, * = $p \leq .05$; ** = $p \leq .01$; - = best estimation of α in opposite direction; 0 = bad fit of regression model; $(-\gamma)$ = skeletal age left out.

lichamelijke activiteit en/of biologische leeftijd op deze factor. Uit de toetsen 2 en 3 bleek echter dat het effect niet kon worden aangetoond voor de onafhankelijke variabele en/of buitenschoolse lichamelijke activiteit. Het is mogelijk dat ons onbekende, niet in het covariantieanalysemodel opgenomen variabelen het effect op de factor spier beter en/of zinvoller zouden kunnen verklaren. Deze gedachte vindt steun in de omstandigheid dat bij vijf van de twaalf hypothesen is gebleken dat de verklarende kracht van het gekozen covariantieanalysemodel gering was (in tabel 8.1. aangegeven met 0), hetgeen verder nog blijkt uit het feit dat slechts in twee van de twaalf gevallen de nulhypothese onder toets 1 kon worden verworpen. Bij de factoren ventilatie, duurkracht en extraversie bleek bovendien dat de beste schatting van α in model 1 (zie 6.3.) tegengesteld was aan het voorspelde teken (in tabel 8.1. aangegeven met -).

Conclusie: alléén op de factor explosieve armkracht kon een significante invloed van de onafhankelijke variabele worden aangetoond ten gunste van de experimentele groep.

8.2. RESULTATEN NIET GEREDUCEERDE MULTIVARIATE ANALYSE

Bij de toetsing van de in hoofdstuk 7.4. opgestelde hypothesen door middel van een ongereduceerde multivariate analyse (zie 6.4.), is getracht een effect aan te tonen van de onafhankelijke variabele. Twee hypothesen werden getoetst:

- De eerste hypothese geldt ten aanzien van de verschilscores op de tien concepten van de semantische differentiaal bij de experimentele groep ten opzichte van de controlegroep. Het ging hierbij om de volgende concepten: leren, doorzettingsvermogen, samenwerken, vroeg opstaan, hobby's, transpireren, sport, elkaar helpen, rekenen en inspanning. De nulhypothese dat deze verschilscores gelijk zijn bij beide groepen kon niet worden verworpen ($p = 0{,}32$). Dit betekent dat geen verschil in attitude tussen beide groepen kon worden aangetoond. Er zij op gewezen dat de covariantiematrices bij beide groepen niet teveel van elkaar verschilden, zodat de toetsing als gerechtvaardigd kan worden beschouwd.

De tweede hypothese geldt ten aanzien van de verschilscores op de variabelen SAGS O, SAGS G, PMT-K P, PMT-K F+, PMT-K F-, gymnastiek voorkeur en gymnastiek rangnummer bij de experimentele groep ten opzichte van de verschilscores op deze variabelen bij de controlegroep. De nulhypothese dat deze verschilscores bij beide groepen gelijk zijn, kon worden verworpen ($p = 0,02$). Ook hier bleken de covariantiematrices niet teveel van elkaar te verschillen zodat de toetsing als gerechtvaardigd kan worden beschouwd. Echter kon niet worden nagegaan welke van de variabele(n) het verschil veroorzaakte(n), omdat in alle per variabele berekende betrouwbaarheidsintervallen het getal nul voorkwam. Dit zou kunnen betekenen dat alle variabelen elk slechts een niet aantoonbaar klein beetje verschillen, terwijl het totale verschil voor alle tezamen wel significant bleek te zijn. In hoofdstuk 9.2. zal hierop nader worden ingegaan.

8.3. RESULTATEN GYMNOLOGISCHE VARIABELEN

Tijdens vier periodes van het schooljaar werden in totaal 31 gymnologische tests afgenomen ter evaluatie van het onderwijs in de lichamelijke oefening.

Alle gymnologische tests waren zo samengesteld dat het klasgemiddelde tussen de twee en drie zou liggen. In werkelijkheid bleek dit goed overeen te komen met een gemiddelde score tussen 2,0 en 2,6. Van deze objectieve prestatietest werden per periode de scores opgeteld en vervolgens per groep gemiddeld.

In tabel 8.2. zijn de gemiddelden en standaardafwijking van de gemiddelde scores over de vier tijdstippen vermeld voor de experimentele en de controlegroep.

Over deze gegevens werd een analyse uitgevoerd (zie 6.6.); daaruit bleek dat de experimentele groep ten opzichte van de controlegroep over de vier meettijdstippen een significant hogere score ($p \leq 0,009$) behaalde.

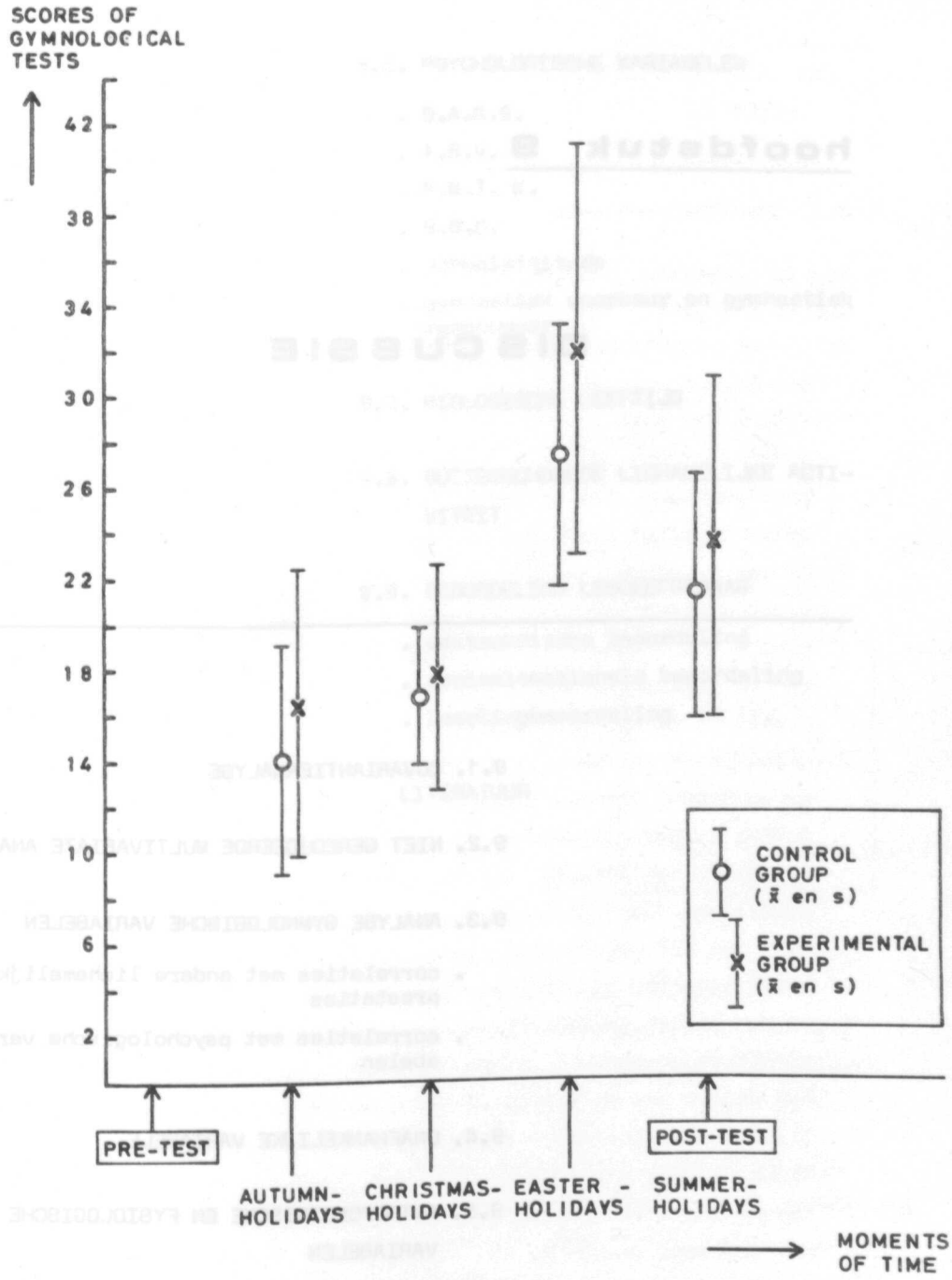
In figuur 8.3. zijn de gegevens van tabel 8.2. grafisch uitgezet. Hoewel de analyse een significante verbetering van de experimentele groep ten opzichte van de controlegroep aangeeft, valt op dat de

scores of gymnological tests		autumn	Christmas	Easter	summer
experim. group	mean	16.12	17.79	31.94	23.62
	s.d.	6.37	5.88	9.11	7.55
control group	mean	14.05	16.97	27.30	21.11
	s.d.	5.05	3.66	6.12	5.56

Tabel 8.2. : Mean and standard deviation of the scores of gymnological tests of experimental and control group during four periods of the school year.

standaardafwijkingen per periode zeer groot zijn; dit is het gevolg van de aanzienlijke prestatieverschillen binnen de groep.

De controlegroep vertoont op elk tijdstip een kleinere standaardafwijking dan de experimentele groep. In dat opzicht vertoonde de controlegroep een grotere homogeniteit dan de experimentele group.



Figuur 8.3. : Mean and standard deviation of gymnological scores of experimental and control group (ordinate) and periods of measurement (abscissa).

hoofdstuk 9

DISCUSSIE

9.1. COVARIANTIEANALYSE

9.2. NIET GEREDUCEERDE MULTIVARIATE ANALYSE

9.3. ANALYSE GYMNOLOGISCHE VARIABELEN

- correlaties met andere lichamelijke prestaties
- correlaties met psychologische variabelen

9.4. ONAFHANKELIJKE VARIABELE

9.5. ANTROPOMETRISCHE EN FYSIOLOGISCHE VARIABELEN

- antropometrische variabelen
- lichamelijk prestatievermogen
- W_{170}

9.6. PSYCHOLOGISCHE VARIABELEN

- S.A.G.S.
- A.B.V.
- P.M.T. K.
- S.S.D.
- schoolattitude
- gymnastiek voorkeur en gymnastiek rangnummer

9.7. BIOLOGISCHE LEEFTIJD

9.8. BUITENSCHOOLSE LICHAMELIJKE ACTIVITEIT

9.9. BEOORDELING LESGEEFGEDRAG

- vaktechnische beoordeling
- sociaalemotionele beoordeling
- leerlingbeoordeling

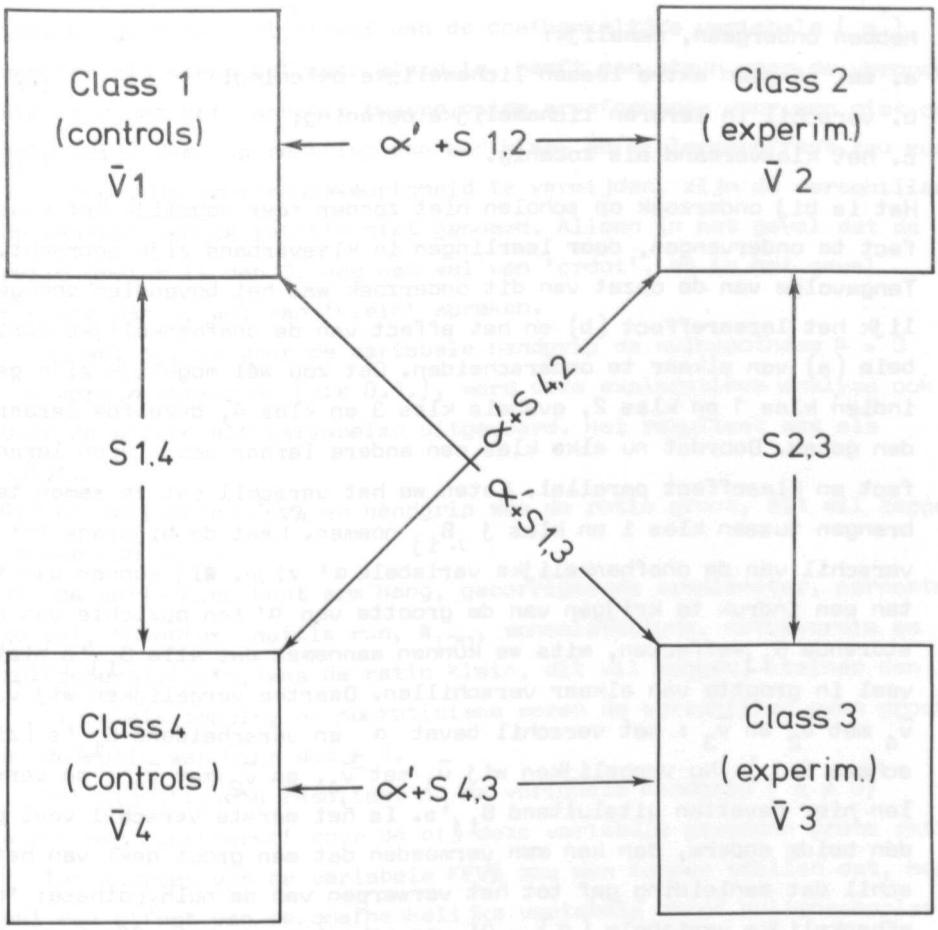
LITERATUUR

9.1. COVARIANTIEANALYSE

Bij de gebruikte statistische analysemethode passen de volgende kanttekeningen:

1. in het experiment waren met opzet veel variabelen opgenomen om vervolgens met behulp van factoranalyse te trachten tot een drastische reductie van het aantal variabelen te komen. Opgemerkt moet worden dat bij sommige van de factoranalyses de ratio variabelen/proefpersonen niet optimaal was, hetgeen mogelijkerwijs de resultaten negatief heeft beïnvloed;
2. er kan gesproken worden van een 'designfout' in de experimentele opzet doordat het lesgeefgedrag van de leraren lichamelijke oefening - een belangrijke storende variabele in het onderzoek - niet aan de leerlingen zelf is gemeten, waardoor deze variabele niet direct in de covariantieanalyse kon worden opgenomen;
3. omdat in de klassen van de experimentele en de controlegroep verscheidene leraren hebben lesgegeven in het vak lichamelijke oefening is het uitgesloten een gevonden verschil tussen deze groepen eenduidig aan effecten van de onafhankelijke variabele toe te schrijven, daar immers het verschil tussen de leraren debet kan zijn aan dit (eventuele) verschil. Slechts door kunstgrepen achteraf kon hier nog iets aan worden gedaan;
4. in een aantal gevallen bleek het toegepaste covariantieanalysemodel slecht te passen bij de waarnemingen (zie 8.1.). Dit zou erop kunnen wijzen dat de gebruikte verklarende variabelen - de onafhankelijke en de storende - zo ze al enige invloed hebben op de afhankelijke variabelen, toch slechts van ondergeschikte betekenis zijn ten opzichte van andere, ons onbekende factoren, zulks in strijd met wat werd verwacht.

ad 3. Over de eventuele invloed van een klas en/of leraareffect op de resultaten kan men, op exploratieve wijze tewerkgaand, het volgende stellen:



Schema 9.1. : Differences between classes:
α' = the effect of the independant variable,
$\bar{v}_1$, $\bar{v}_2$, $\bar{v}_3$ and $\bar{v}_4$ mean difference-scores of the four classes; s = the effect of teacher and/or class.

Beschouw de gemiddelde verschilscores op een variabele in de vier klassen, $\bar{v}_1$, $\bar{v}_2$, $\bar{v}_3$ en $\bar{v}_4$. Kan een verschil tussen $\bar{v}_1$ en $\bar{v}_2$ worden toegeschreven aan de experimentele conditie? Door het middelen over personen mag men aannemen dat persoonseffecten zijn uitgeschakeld nadat voor verschillen in buitenschoolse activiteit en verschillen in biologische leeftijd is gecorrigeerd. Men kan de opgetreden verschillen dan terugvoeren op de verschillende 'behandeling' die de klassen

hebben ondergaan, namelijk:

a. wel of geen extra lessen lichamelijke oefening;

b. verschil in leraren lichamelijke oefening;

c. het klasverband als zodanig.

Het is bij onderzoek op scholen niet zonder meer mogelijk het klaseffect te ondervangen, daar leerlingen in klasverband zijn gebracht. Tengevolge van de opzet van dit onderzoek was het bovendien onmogelijk het leraareffect (b) en het effect van de onafhankelijke variabele (a) van elkaar te onderscheiden. Dat zou wél mogelijk zijn geweest indien klas 1 en klas 2, evenals klas 3 en klas 4, dezelfde leraar hadden gehad. Doordat nu elke klas een andere leraar had, lopen leraareffect en klaseffect parallel. Laten we het verschil dat ze samen teweegbrengen tussen klas i en klas j S_{ij} noemen. Laat de bijdrage tot het verschil van de onafhankelijke variabele α' zijn. Wij kunnen dan trachten een indruk te krijgen van de grootte van α' ten opzichte van de storende S_{ij}-effecten, mits we kunnen aannemen dat alle S_{ij}'s niet teveel in grootte van elkaar verschillen. Daartoe vergelijken wij $\bar{v}_1$ en $\bar{v}_4$ met $\bar{v}_2$ en $\bar{v}_3$: het verschil bevat α' en verscheidene S_{ij}'s (zie schema 9.1.). Nu vergelijken wij $\bar{v}_1$ met $\bar{v}_4$, en $\bar{v}_2$ met $\bar{v}_3$: de verschillen hier bevatten uitsluitend S_{ij}'s. Is het eerste verschil veel groter dan beide andere, dan kan men vermoeden dat een groot deel van het verschil dat aanleiding gaf tot het verwerpen van de nulhypothese: 'de onafhankelijke variabele (α) = 0' op rekening van het effect van de onafhankelijke variabele α' komt.

De hierboven beschreven exploratieve analysemethode is uitgevoerd over de variabelen die werden opgenomen in de covariantieanalyse (zie 8.1.). In feite is dus, per variabele de volgende ratio berekend:

$$\frac{\text{effect v.d.onafh.var.+ klas- en/of leraareff. tussen exp. en contr.groep}}{\text{klas-en/of leraareff. binnen exp. en contr. groep}}$$

of, anders geformuleerd:

$$\frac{(\bar{v}_2 + \bar{v}_3) - (\bar{v}_1 + \bar{v}_4)}{(\bar{v}_2 - \bar{v}_3) + (\bar{v}_1 - \bar{v}_4)}$$

(zie schema 9.1.). Indien deze ratio groot is (> 3), kan men het eventuele verschil tussen experimentele en controlegroep vermoedelijk

toeschrijven aan het effect van de onafhankelijke variabele (α'), wanneer hij klein tot zeer klein is, heeft men steun voor de veronderstelling dat het verschil tussen beide proefgroepen voor een niet onbelangrijk deel op rekening van een klas- en/of leraareffect zou kunnen komen. Om schijnnauwkeurigheid te vermijden, zijn de verschillende waarden van de ratio's niet genoemd. Alleen in het geval dat de ratio groter is dan 3, mag men wel van 'groot', en in het geval kleiner dan 1, wel van 'klein' spreken.

Hoewel alleen voor de variabele handgrip de nulhypothese $\alpha = 0$ kon worden verworpen (zie 8.1.), werd deze exploratieve analyse ook over de andere elf variabelen uitgevoerd. Het resultaat was als volgt:
Bij de variabelen FEV% en handgrip was de ratio groot, dit wil zeggen, groter dan 3.
Bij de variabelen bent arm hang, gecorrigeerde armdiameter, percentage vet, 50 meter shuttle run, W_{170}, schoolattitude, extraversie en zelfdefensiviteit, was de ratio klein, dit wil zeggen, kleiner dan 1.

Bij plate tapping en neuroticisme waren de verschillen even groot en de ratio was hier dus $\pm$ 1.

Het significante resultaat op de variabele handgrip ($\alpha \neq 0$) wordt nog 'versterkt' door de bij deze variabele gevonden grote ratio.

Ten aanzien van de variabele FEV% zou men kunnen stellen dat, hoewel een effect van de onafhankelijke variabele en/of buitenschoolse lichamelijke activiteit en/of biologische leeftijd niet kon worden aangetoond, deze hierop wellicht meer invloed hebben gehad dan een klas- en/of leraareffect.

Bij de variabelen waar de ratio klein was, kan men wellicht opmerken dat er een kleine indicatie bestaat dat klas- en/of leraareffecten een belangrijke rol hebben gespeeld. Dit geldt dan in iets sterkere mate voor de bent arm hang, waar de ratio zeer klein bleek te zijn.

Op grond van het bovenstaande zou men samenvattend kunnen stellen dat het enige significante resultaat van de covariantieanalyse, namelijk het aangetoonde verschil op de variabele handgrip, nog wordt versterkt door de resultaten van de exploratieve analyse: het verschil tussen beide proefgroepen op deze variabele komt geheel op

rekening van de onafhankelijke variabele.

Ten aanzien van een groot deel van de overige in de covariantie-analyse opgenomen variabelen is het wellicht het meest zinnig om op te merken dat de resultaten van de exploratieve analyse aanleiding geven tot het niet uitsluiten van de mogelijkheid dat in een aantal gevallen het eventuele effect van de onafhankelijke variabele is overschaduwd of misschien tegengewerkt door verschillen tussen leraren er/of klassen.

In dit onderzoek is wel een meting opgenomen over de sociometrische structuur binnen de klassen, die wellicht iets zou kunnen zeggen over een eventueel klaseffect (zie 9.6.), hoewel hierop geen significant resultaat werd gevonden (zie 8.2.). Omtrent eventuele verschillen tussen leraren, dus een 'leraareffect' geven de resultaten van de vaktechnische (zie 5.3.1.) en de sociaalemotionele (zie 5.3.2.) beoordeling van het lesgeefgedrag van de leraren lichamelijke oefening aanleiding met een dergelijk effect rekening te houden (zie 9.9.).

ad 4. Zoals reeds in hoofdstuk 8.1. werd gesteld, kon in tien van de twaalf gevallen de nulhypothese $\alpha = \beta^{(e)} = \beta^{(c)} = \gamma = 0$ niet worden verworpen. Dit betekent in feite dat het door ons gekozen model geen grotere verklaringskracht had dan een triviaal model zonder deze verklarende variabelen, en in dat geval dus de vorm $\mu + \Omega$ zou hebben, dat wil zeggen, het populatiegemiddeld (μ) plus de errorvariantie (Ω). Bovendien bleek in vijf van deze tien gevallen de regressievergelijking wel opmerkelijk slecht te zijn, dat wil zeggen, dat het percentage verklaarde variantie zeer laag bleek te zijn (in tabel 8.1. aangegeven met 0). Los van de hierboven onder ad 3. genoemde mogelijkheid van klas- en/of leraareffecten, zou dit ook - of mede - kunnen betekenen dat de door ons gebruikte verklarende variabelen (de onafhankelijke en de storende) nauwelijks van invloed zijn geweest op de afhankelijke variabelen, en/of van ondergeschikte betekenis zijn geweest in vergelijking met andere, ons onbekende factoren. Wat betreft de onafhankelijke variabele zou men bij voorbeeld kunnen denken aan de mogelijkheid dat het niveauverschil op deze variabele, drie lessen ten opzichte van vijf lessen lichamelijke oefening, te gering

is geweest om tot verschillen in resultaten van beide proefgroepen te kunnen leiden (zie 9.8.). Ten aanzien van de storende variabelen, buitenschoolse lichamelijke activiteit en biologische leeftijd, kan men zich afvragen of ze een juiste keuze zijn geweest. Het is wellicht onjuist geweest de interactie tussen de onafhankelijke variabele en biologische leeftijd niet in het covariantieanalysemodel op te nemen.

Het blijft een bevreemdend resultaat dat de beste schatting van de onafhankelijke variabele (α) in het door ons gekozen covariantieanalysemodel (zie 6.3., model 1) bij de variabelen FEV%, bent arm hang en extraversie tegengesteld aan het voorspelde teken bleek te zijn, hetgeen betekent dat de opgetreden (niet significante) verschillen tussen beide proefgroepen in de 'verkeerde' richting optraden (zie tabel 8.1.).

Ten aanzien van de factor 'spier', gerepresenteerd door de variabele gecorrigeerde armdiameter, is, bij wijze van uitzondering, nog een nadere analyse uitgevoerd. Toets 1 gaf namelijk aan dat het zinvol was om verder te gaan zoeken naar mogelijke effecten van de onafhankelijke variabele en/of buitenschoolse lichamelijke activiteit en/of biologische leeftijd. Uit de toetsen 2 en 3 bleek echter dat het effect niet kon worden aangetoond voor de onafhankelijke variabele en/of buitenschoolse lichamelijke activiteit. Daarom is in dit speciale geval nader geanalyseerd of wellicht de biologische leeftijd dit effect zou kunnen verklaren. Deze nadere analyse wijst erop dat, binnen het door ons gekozen model, een effect van de biologische leeftijd zou kunnen worden aangetoond op de variabele gecorrigeerde armdiameter, zij het met een 'significantie' van $p < 0,10$. Ook hier is het natuurlijk zeer wel mogelijk dat andere variabelen het effect op de gecorrigeerde armdiameter beter, significanter en/of zinvoller zouden kunnen verklaren.

9.2. NIET GEREDUCEERDE MULTIVARIATE ANALYSE

Voor de groepjes waarbij de factoranalyse geen gelijkluidende structuur in experimentele en controlegroep opleverde, was een verantwoorde reductie tot een beperkt aantal representerende variabelen

niet mogelijk. In dergelijke gevallen werd besloten tot een ongereduceerd oplossen van het tweesteekproevenprobleem. De variabelen die tot deze groepjes behoorden, werden met behulp van Hotelling's T^2 onderzocht, waarbij expliciet rekening werd gehouden met de mogelijkheid van onderlinge afhankelijkheid van deze variabelen. Univariate, onafhankelijke toepassingen van de t-toets waren niet mogelijk (zie 6.5.). Ingaande op de onder hoofdstuk 8.2. vermelde resultaten van deze multivariate analyse kan het volgende worden opgemerkt.

Ten aanzien van de eerste groep variabelen, de attitudes ten opzichte van de 10 SSD-concepten, kon geen verschil worden aangetoond tussen beide proefgroepen. Omdat Hotelling's T^2 echter altijd tweezijdig is, is het interessant om na te gaan of de verschillen op de 10 concepten als gunstig voor de experimentele groep kunnen worden beschouwd. Bij 9 van de 10 concepten bleek dit inderdaad het geval. Beide proefgroepen vertoonden weliswaar een afname in attitudescore - met uitzondering van het concept sport - ten aanzien van de concepten, maar deze afname is altijd kleiner bij de experimentele groep. De verschillen zijn echter van die aard dat zelfs in het hypothetische geval van 10 gerechtvaardigde univariate toepassingen van de t-toets geen van deze verschillen de significantiedrempel zou hebben overschreden. Zeker in het geval van de attitude ten aanzien van de concepten hobby's en sport is dit resultaat afwijkend van het vooronderzoek waarbij verandering op deze concepten ten gunste van de experimentele groep werden gevonden.

Ten aanzien van de tweede groep variabelen, te weten, SAGS (O en G), PMT-K (P, F+ en F-) en gymnastiek (gymnastiek voorkeur en rangnummer) die multivariaat werden geanalyseerd, werd gevonden dat deze groep variabelen als geheel significant verschil aangaf tussen experimentele en controlegroep. Omdat alle betrouwbaarheidsintervallen het getal nul bevatten, kon echter niet worden nagegaan welke variabele(n) het verschil veroorzaakte(n). Bij nadere analyse van de verschilscores op deze variabelen lijkt het echter dat vooral de variabelen SAGS O en SAGS G het verschil tussen beide proefgroepen veroorzaken. Dit nu zijn juist de variabelen die iets zeggen over het groepsverband, met andere woorden: de mogelijkheid van een

'klaseffect' wordt door deze bevinding enigszins plausibel (zie 9.1. en 9.6.). Rest nog te vermelden dat de 'verschillen' op deze twee sociometrische maten als gunstig voor de experimentele groep kunnen worden omschreven.

9.3. ANALYSE GYMNOLOGISCHE VARIABELEN

Het feit dat de hypothese over een grotere toename van de gymnologische prestaties bij de experimentele groep ten opzichte van de controlegroep bevestigd kon worden, geeft op dit punt steun aan de resultaten van het vooronderzoek (Oud e.a., 1971).

Terwijl in dit vooronderzoek alle gymnologische tests in beide groepen door dezelfde leraar werden afgenomen, geeft dit resultaat, waarbij door vier verschillende leraren deze zelfde tests werden afgenomen, aanleiding te veronderstellen dat deze tests op een redelijke wijze de vorderingen in de lichamelijke oefening objectiveerbaar hebben gemaakt.

Bij nadere toetsing van het verschil in gymnastiekprestaties op lineairiteit bleek dit niet te kunnen worden bevestigd: uit de verschilscores op de vier tijdstippen in figuur 8.3. is dit ook duidelijk te zien. Een dergelijke lineaire trend zou men eigenlijk ook niet verwachten wanneer het gaat om trainingseffecten; bij gelijke trainingsintensiteit en trainingsduur, maar met een vaste verhoogde frequentie (drie lessen om vijf lessen per week), moet worden verwacht dat de grootte van het trainingseffect steeds geringer wordt.

Wanneer men de totaal behaalde score op alle 31 tests correleert met de scores behaald over elke periode, blijkt dat de correlaties bij zowel experimentele als controlegroep zeer hoog zijn (zie tabel 9.2.).

Correlaties met andere lichamelijke prestaties

Omdat deze tests tot dusver op geen enkele wijze gevalideerd zijn, is het interessant om de prestaties, behaald op deze gymnologische tests te vergelijken met de eenvoudige prestatietests, waarvan wel de factorstructuren bekend zijn.

	Total score of gymn. tests during four periods	
score of gymn. tests per period	experimental group	control group
autumn	+ .88	+ .91
Christmas	+ .93	+ .75
Easter	+ .91	+ .81
summer	+ .94	+ .82

Tabel 9.2. : Spearman correlations between total score and score of each period of experimental and control group.

In tabel 9.3. zijn de Spearman correlatiecoëfficiënten vermeld van de gemiddelde somscore van alle eenendertig gymnologische tests met de scores van de eenvoudige prestatietests op de voortest. Bij de experimentele groep zijn alle correlaties significant, bij de controlegroep eveneens met uitzondering van de handgrip (+ 0,23).

Uit deze correlaties zou men kunnen besluiten dat de verschillende grondeigenschappen,die verondersteld worden door de eenvoudige prestatietests te worden gemeten, redelijk vertegenwoordigd zijn in de gymnologische tests.

Correlaties met psychologische variabelen

Zoals reeds eerder werd opgemerkt, bleken op de voortest een aantal opvallende verschillen te bestaan in factorstructuur van de psychologische variabelen tussen experimentele en controlegroep (zie 7.2.). Deze verschillen komen ook tot uiting bij het zoeken naar verbanden tussen de gymnologische tests en psychologische variabelen. In tabel 9.4. zijn alleen die correlatiecoëfficiënten vermeld die significant waren op 5% niveau, zowel bij de experimentele als de controlegroep. Duidelijk blijkt dat een hoge gymnastiekvoorkeur samengaat met hoge prestaties in de gymnastiekles. Jongens met een hoge extraversiescore zijn ook beter in gymnastiek. De hypothese, dat de onafhankelijke variabele effect heeft op de E-score (verhoging) kon echter

simple performance tests on pretest	total score of gymn. tests	
	experimental group	control group
plate tapping	- .36*	- .29*
sit and reach	+ .69**	+ .38**
50 meter shuttle run	- .63**	- .39**
bent arm hang	+ .57**	+ .32*
vertical jump	+ .56**	+ .49**
handgrip	+ .46**	+ .23
12 min. run walk	+ .57**	+ .28*

Tabel 9.3. : Spearman correlations between simple performance tests on pretest, and total score of all gymnological tests during the schoolyear (* = $p \leq .05$; ** = $p \leq .01$).

psychological variables	total score of gymn. tests	
	experimental group	control group
preference of physical education	+ .35*	+ .40**
ABV- extraversion	+ .42**	+ .35*

Tabel 9.4. : Spearman correlations between total score of gymnological tests and psychological variables of experimental and control group; only correlations are given that were significant in both groups (* = $p \leq .05$; ** = $p \leq .01$).

niet worden verworpen.

9.4. ONAFHANKELIJKE VARIABELEN

Omdat jaarplan, leerplan, testplan van de brugklassen reeds een

aantal jaren voorafgaand aan het experiment waren beproefd door de leraren lichamelijke oefening, leverden de twee wekelijkse lessen die besteed werden aan gymnastiek, spel en athletiek onderwijs, bij de uitvoering van het lessenplan weinig problemen op. De vorderingen van de vier klassen liepen niet ver uiteen zodat er nooit aanleiding bestond om bij een van de klassen duidelijk en definitief af te wijken van het opgestelde jaarplan.

Het lesplan voor de wekelijkse zwemles werd week voor week, na beëindiging van de voorgaande zwemles vastgesteld; in gezamenlijk overleg werd, afhankelijk van de gemaakte vorderingen, voor elke prestatiegroep de inhoud vastgesteld. Ook mutaties in de samenstelling van de prestatiegroepen werden tegelijkertijd besproken.

Met betrekking tot de extra lessen werd na onderling overleg met de twee betreffende leraren de volgorde van deze lessen bepaald. Hoewel ook hierbij in principe vooraf een volgorde was gemaakt, bleek het dikwijls nodig hiervan af te wijken. Dit kwam doordat deze lessen betrekkelijk nieuw waren en niet eerder op grote schaal waren beproefd.

De lessen lichamelijke oefening werden voor het grootste deel gegeven in twee boven elkaar gelegen gymnastiekzalen. Afmetingen van deze zalen waren gelijk, de inrichting ervan verschilde. De verschillen betroffen zowel de 'vaste' toestellen (zoals ringen, touwen, rekstok, wandrek, ladders) als de verplaatsbare toestellen (zoals bruggen, kasten, paarden, bokken, lange mat). Deze laatste waren zoveel mogelijk in één zaal gegroepeerd.

Elke week werd door de vier leraren lichamelijke oefening van zaal gewisseld.Hierdoor konden de opgestelde lessen door elke klas niet in precies dezelfde volgorde worden gegeven; in de praktijk bleek een spreiding van hoogstens veertien dagen op te treden.

Door incidenteel onderling contact van leerlingen uit de experimentele groep en controlegroep werd een enkele maal aan de leraren lichamelijke oefening de vraag gesteld waarom aan de ene klas meer lessen lichamelijke oefening werden gegeven dan aan de andere. Als reden werd aangevoerd dat voor die bepaalde klas op het rooster een aantal lesuren 'over' waren.

9.5. ANTROPOMETRISCHE EN FYSIOLOGISCHE VARIABELEN

Antropometrische variabelen

De antropometrische variabelen werden in dit onderzoek opgenomen om de bouw en samenstelling van het lichaam te meten. De gegevens van de voortest zijn weergegeven in tabel 9.5..

pre-test data of anthropometric variables

measurement	unit	group	mean	S.D.	minimum	maximum	range
weight	kg	contr.	39.9	6.6	28.7	54.0	25.3
		exp.	42.2	5.9	30.2	57.5	27.3
height	cm	contr.	154.5	7.8	139.2	178.1	38.9
		exp.	155.5	7.0	137.0	166.9	29.9
biacromial diameter	cm	contr.	32.4	1.8	29.0	36.7	7.7
		exp.	31.9	2.2	25.6	35.9	19.3
biiliocristal diameter	cm	contr.	23.1	1.2	20.8	25.4	4.6
		exp.	23.2	1.2	19.2	25.2	6.0
upper arm circumference	cm	contr.	19.7	1.9	16.1	24.3	8.2
		exp.	20.4	1.8	17.2	25.4	8.2
thigh circumference	cm	contr.	42.4	4.3	34.9	53.5	18.6
		exp.	43.6	4.2	35.3	53.9	18.6
calf circumference	cm	contr.	29.5	2.1	25.4	34.6	9.2
		exp.	30.3	2.0	26.3	34.6	8.3
wrist breadth	cm	contr.	4.85	0.34	4.27	5.74	1.47
		exp.	4.96	0.28	4.44	5.62	1.18
bicondylar femur	cm	contr.	9.06	0.45	8.40	10.05	1.65
		exp.	9.17	0.42	8.17	10.10	1.93
bicipital skinfold	mm	contr.	5.4	2.2	2.7	12.6	9.9
		exp.	5.9	2.8	2.8	14.4	11.6
tricipital skinfold	mm	contr.	10.8	3.5	5.2	19.8	14.6
		exp.	11.8	4.7	4.6	25.8	21.2
subscapular skinfold	mm	contr.	6.8	2.5	4.0	17.0	13.0
		exp.	7.2	2.3	4.4	14.4	10.0
suprailiacal skinfold	mm	contr.	7.8	3.5	3.8	20.0	16.2
		exp.	9.1	4.5	4.0	24.6	20.6
corrected upper arm diameter	cm	contr.	5.2	0.4	4.3	6.3	2.0
		exp.	5.3	0.3	4.8	6.2	1.4
% fat	%	contr.	17.5	3.4	10.1	26.7	16.6
		exp.	18.2	3.8	11.7	27.3	15.6

Tabel 9.5. : Pre-test scores of anthropometric variables; mean, standard deviation, minimum, maximum and range are given for the control (n = 37) and the experimental (n = 33) group.

De verwachting dat de spiermassa meer zou toenemen bij de experimentele dan bij de controlegroep, en dat het percentage vet juist meer zou afnemen, kon niet worden bevestigd. Hoewel de veranderingen in

de verwachte richting plaatsvonden, en de gemiddelden van experimentele en controlegroep verschilden op een wijze zoals voorspeld, bleken de verschillen tussen beide groepen niet significant te zijn (zie tabel 9.6.).

pre- and post-test data of percentage of fat and muscle mass						
measurement	group	pretest mean	pretest S.D.	posttest mean	posttest S.D.	difference
percentage of fat	contr.	17.5	3.4	16.4	3.2	- 1.1
	exp.	18.2	3.8	16.7	3.4	- 1.5
muscle mass	contr.	5.2	0.4	5.8	4.9	+ 0.6
	exp.	5.3	0.3	5.9	4.6	+ 0.6

Tabel 9.6. : Pre- and post-test scores of percentage of fat and corrected upper arm diameter; mean, standard deviation and difference of scores of experimental (n = 33) and control group (n = 37).

De vermindering in onderhuids vet is in overeenstemming met de bevindingen van De Wijn e.a. (1968). Zij vonden bij dezelfde leeftijdsgroep na één jaar groei een afname bij zeven van de tien gemeten huidplooien, waaronder drie van de vier huidplooien die in het onderhavige onderzoek werden gemeten.

Een toename van de dagelijkse hoeveelheid lichamelijke activiteit zou tevens van invloed hebben kunnen zijn. De overgang van lager naar voortgezet onderwijs betekent namelijk dat de leerlingen over het algemeen een grotere afstand tussen huis en school moeten overbruggen. Aangezien deze vorm van activiteit 30,8% van het totaal uitmaakte (zie 5.2.), kan deze verandering van grote invloed zijn geweest. Verder vindt veelal een toename plaats van de activiteiten in verenigingsverband, en met name in sportverenigingen.

Ten slotte kan ook de toename van het aantal lessen lichamelijke oefening per week tot drie - ten opzichte van het aantal lessen op de lagere school met maximaal twee - een bijdrage aan de toename van de dagelijkse hoeveelheid lichamelijke activiteit betekenen.

Zeker wanneer we in aanmerking nemen dat de dagelijkse lichamelijke activiteit is toegenomen, is het ontbreken van significante verschillen

tussen experimentele en controlegroep mogelijk veroorzaakt door het feit dat de twee extra lessen lichamelijke oefening per week een te geringe vergroting van de normale dagelijkse lichamelijke activiteit betekende om verschillen in adaptatie van lichaamsbouw en -samenstelling te kunnen krijgen. Gezien de tamelijk hoge correlatie tussen spiermassa, bepaald door de gecorrigeerde armdiameter, en de skeletleeftijd (r = 0,45 en 0,53 bij respectievelijk experimentele en controlegroep) is de invloed van ontwikkelingsfactoren op de hoeveelheid spiermassa bij deze 12 en 13-jarige jongens waarschijnlijk groot. Dit wordt bevestigd door de resultaten van de covariantieanalyse waar een zeker effect van de biologische leeftijd op de spiermassa kon worden aangetoond (9.1.). Volgens Klissouras (1973) bestaat de mogelijkheid dat bij kinderen in de groeispurt de hormonale ontwikkeling optimaal is voor groei en ontwikkeling zodat extra trainingsprikkels geen waarneembaar effect meer hebben.

Lichamelijk prestatievermogen

Voor het meten van het lichamelijk prestatievermogen werd onder andere gebruikgemaakt van lichamelijke prestatietests. De voortestgegevens van zowel experimentele als controlegroep zijn weergegeven in tabel 9.7.

De lichamelijke prestatietests waren zes tests voor het meten van prestaties van het neuromusculaire systeem, één test voor het meten van de ventilatieprestatie, en één voor het aerobe vermogen. Deze laatste, de 12-minutenloop, kwam te vervallen omdat hij nergens een aparte factor vormde en de correlatie met de W_{170} (r =+0,37), een betrouwbare maat voor het aerobe vermogen, zelfs lager was dan met een aantal andere tests, zoals de 50 meter shuttle run (r =-0,61), de bent arm hang (r =+0,52) en de vertical jump (r =+0,38).

Cumming (1973) meldde dat de reproduceerbaarheid van de 12-minutenloop aanmerkelijk hoger was bij een groep kinderen van 12 tot 17 jaar die goed gemotiveerd waren, dan bij een groep vrijwel even oude, niet geselecteerde schoolkinderen. Gesteld kan daarom worden dat, hoewel de 12-minutenloop een goede voorspeller van het aerobe vermogen kan zijn, de test in dit onderzoek vermoedelijk meer als een maat voor de motivatie heeft gefungeerd.

Overigens moet erop worden gewezen dat vrijwel alle tests met

pre-test data of physical performance tests							
measurement	unit	group	mean	S.D.	minimum	maximum	range
plate tapping	sec	contr.	20.9	2.4	17.4	30.0	12.6
		exp.	20.4	2.1	16.8	24.6	7.8
sit and reach	cm	contr.	27.9	6.0	14.0	37.0	23.0
		exp.	26.8	5.9	11.0	35.0	24.0
50 m. shuttle run	sec	contr.	14.1	0.6	12.3	15.3	3.0
		exp.	14.2	0.8	12.8	15.7	2,9
bent arm hang	sec	contr.	18.4	9.4	6.0	41.0	35.0
		exp.	21.8	11.4	4.0	48.0	44.0
vertical jump	cm	contr.	29.1	5.1	17.0	46.0	29.0
		exp.	29.0	4.9	21.0	39.0	18.0
handgrip	kg	contr.	21.5	5.1	12.0	33.0	21.0
		exp.	22.1	4.5	14.0	32.0	18.0
12 min. run walk	m	contr.	2388	235	1848	2808	960
		exp.	2400	299	1488	2832	1344
W_{170}	watt/kg	contr.	2.69	0.56	1.78	3.83	2.05
		exp.	2.68	0.58	1.65	4.02	2.37
FEV %	%	contr.	86.9	6.2	72.4	97.4	25.0
		exp.	87.4	6.4	72.0	96.6	24.6

Tabel 9.7. : Pre-test scores of the physical performance tests. Mean, standard deviation, minimum, maximum and range are given for the control (n = 37) and the experimental group (n = 33).

elkaar correleerden en wel zodanig dat een goede prestatie op de ene test samenging met een redelijk goede prestatie op andere tests (zie tabel 9.8.).

Dit zou erop kunnen wijzen dat de eigenschappen die door de verschillende tests worden gemeten, bij 12 en 13-jarige jongens niet zo duidelijk te onderscheiden zijn.

Willems e.a. (1971) onderzochten de aanwezigheid van verschillende krachtfactoren bij 12 tot 19-jarigen, daarbij rekening houdend met lichaamsgewicht en lichaamsbouw van de proefpersonen. De resultaten geven aan dat het onderscheid in verschillende krachtfactoren tijdens de groeiperiode toeneemt naarmate de leeftijd stijgt. Men concludeerde dat de uiteenlopende scores op krachttests bij jongeren van 12 tot 14 jaar meer door lichaamsbouw en lichaamsgewicht worden veroorzaakt dan dat bij de ouderen het geval is. De grote

verschillen in lichamelijke ontwikkeling op deze jongere leeftijd brengen grote verschillen in bouw met zich mee waardoor de prestaties zo sterk uiteenlopen.

intercorrelations of pre-test data of physical performance tests

test	1	2	3	4	5	6	7	8	9
plate tapping (1)	X	-.13	+.22*	-.29**	-.01	-.20*	-.13	+.03	-.17
sit and reach (2)		X	-.30**	+.28**	+.33**	+.28**	+.23*	-.05	-.08
50 m. shuttle run (3)			X	-.37**	-.48**	-.23*	-.61**	-.24*	-.18
bent arm hang (4)				X	+.32**	+.24*	+.52**	+.15	-.02
vertical jump (5)					X	+.20*	+.38**	+.15	+.04
handgrip (6)						X	+.10	+.13	+.00
12 min. run walk (7)							X	+.37**	+.16
W_{170} (8)								X	+.14
FEV % (9)									X

Tabel 9.8. : Intercorrelations (Spearman) of pre-test scores of physical performance tests (n = 70);
(* = $p \leq .05$; ** = $p \leq .01$).

Het is dus mogelijk dat aanleg en ontwikkeling op deze leeftijd een zo belangrijke invloed uitoefenen op de prestaties dat het onderscheiden van eigenschappen die het prestatievermogen bepalen vrijwel niet mogelijk is.

Uit de resultaten van de covariantie blijkt dat de explosieve armkracht (handgrip) significant toeneemt bij de experimentele groep ten opzichte van de controlegroep als gevolg van de extra lessen lichamelijke oefening. De maat voor spiermassa (gecorrigeerde armdiameter) neemt toe bij de experimentele groep ten opzichte van de controlegroep ($p < 0,10$) als gevolg van de biologische

leeftijd.

Bij vergelijking van de lichamelijke prestatietests met de variabele gecorrigeerde armdiameter blijkt op de voortest de hoogste correlatie op te treden met handgrip (r = 0,59).

Blijkens deze correlatie is er een verband tussen beide variabelen. De mogelijkheid bestaat dat handgrip behalve door de onafhankelijke variabele tevens is beïnvloed door de storende variabele biologische leeftijd. In de covariantie is echter alleen nagegaan welke invloed de biologische leeftijd in combinatie met de onafhankelijke variabele en de buitenschoolse lichamelijke activiteit heeft op de afhankelijke variabele; er werd niet nagegaan welke de invloed was van de biologische leeftijd alleen.

W_{170}

Het aerobe uithoudingsvermogen, gemeten met de fietsergometerbelasting (W_{170}) bleek bij de experimentele groep niet significant te verbeteren ten opzichte van de controlegroep.

Hoewel in een discussie naar aanleiding van het vooronderzoek (Kemper e.a., 1972, b) werd verondersteld dat door een verfijning van de meetmethoden met behulp van ergometrisch onderzoek eerder een trainingseffect zou kunnen worden aangetoond, bleek dit niet het geval te zijn. Om dit resultaat te interpreteren kan het volgende worden opgemerkt:

De voortest van de fietsproef vond niet vóór het begin van het schooljaar plaats tijdens de introductiedagen, maar in de loop van de maand oktober. De natest vond niet helemaal aan het einde van het schooljaar plaats, maar in de periode tussen half april en half mei. Daarbij was op voor- en natest de volgorde waarin de leerlingen werden getest gelijk. De tussen voor- en natest verlopen periode was evenwel voor elke leerling gelijk, maar aanzienlijk korter dan van de meeste andere variabelen.

Uit de gegevens van tabel 9.9. blijkt dat de W_{170} van gemiddeld 2,7 watt/kg op de voortest is toegenomen tot gemiddeld 3,1 watt/kg op de natest.

Deze toename kan verklaard worden door:

- een afname van het percentage vet met gemiddeld 1%;
- een gewenningseffect waarbij de hartfrequentie bij overigens gelijke belasting op de natest ten opzichte van de voortest daalt als gevolg van een herhaald testen;
- een toegenomen dagelijkse lichamelijke activiteit tijdens de brugklasperiode ten opzichte van de lagere schoolperiode.

De gemiddelde W_{170} is zowel op de voor- als op de natest aanzienlijk hoger dan die welke in 1970 met een vergelijkbare populatie werd gevonden (Kemper, 1972, a) met een gemiddelde waarde van 2,0 watt/kg (± 0,75). Dit verschil zou verklaard kunnen worden uit het feit dat een verschillend belastingschema werd gehanteerd: drie maal zes minuten, in plaats van vier maal drie minuten.

W_{170} (Watt/kg		mean	S.D.	min.	max.	range
experimental group	pretest	2.68	0.58	1.65	4.02	2.37
	posttest	3.10	0.59	1.79	4.46	2.67
control group	pretest	2.69	0.55	1.78	3.83	2.05
	posttest	3.02	0.50	1.91	4.11	2.20

Tabel 9.9. : Mean, standard deviation, minimum, maximum, range of W_{170} (in Watt/kg body weight) in experimental and control group on pre- and posttest.

9.6. PSYCHOLOGISCHE VARIABELEN

S.A.G.S.

Uit de resultaten van de multivariate analyse over groep B, zoals vermeld in hoofdstuk 8.2., blijkt dat er een significant effect is op de afhankelijke variabelen: SAGS O, SAGS G, PMT-K P, PMT-K F-, PMT-K F+, gymnastiek rangnummer en gymnastiek voorkeur. Het was echter niet mogelijk om aan te geven welke afhankelijke

variabele(n) voor het verschil verantwoordelijkheid is(zijn).

test	group	pretest mean	S.D	posttest mean	S.D.	n	diff.	pretest min.max.	posttest min. max.
SAGS O	exp.	2.35	.46	2.50	.58	33	+.15	1.47 3.06	1.33 3.87
	contr.	1.93	.48	1.70	.29	37	-.23	1.30 2.93	1.10 2.27
SAGS G	exp.	2.34	.70	2.48	.70	33	+.14	1.00 3.76	1.00 3.74
	contr.	1.93	.65	1.70	.56	37	-.23	1.05 3.43	1.00 3.73

Tabel 9.10. : Mean, standard deviation and range of the experimental and control group on the SAGS O (accepted) and SAGS G (given), on pretest and posttest; differences between post- and pretest.

Wanneer men de gemiddelden van de experimentele en controlegroep op de voor- en natest ten aanzien van bovengenoemde variabelen vergelijkt, is het misschien mogelijk één of meer variabelen aan te wijzen die verantwoordelijk zijn. Uit de in tabel 9.10. vermelde resultaten blijkt dat alleen het gemiddelde van de ontvangen scores (SAGS O) en het gemiddelde van de gegeven scores (SAGS G) toenemen in de experimentele groep, en afnemen in de controlegroep, met andere woorden, de verschillen in groepscohesie kunnen misschien het verschil veroorzaken. Opvallend is dat in de experimentele en in de controlegroep één klas gelijk blijft in gemiddelde score op de SAGS O en G, maar dat in de experimentele groep één klas een stijging vertoont, en in de controlegroep één klas een daling. Dit zou kunnen wijzen op een mogelijk klas- en/of leraareffect (zie tabel 9.11.).

Omdat besloten werd de gegevens multivariaat te analyseren, is een t-toets achterwege gelaten.

Wat betreft het instrument zelf, twee opmerkingen:

1. Uit de resultaten zijn leerlingen weggelaten (zie 2.1.) op grond van de volgende uitvalcriteria: a. 10% van de lessen lichamelijke oefening gemist; b. afwezigheid van de leerling op een van de tijdstippen van meting. Het gevolg van deze criteria voor de bepaling van de sociometrische structuur binnen de klas is dat de klas als geheel ten onrechte niet intact bleef. Immers, het

is niet aannemelijk dat de sociometrische structuur zich wijzigt onder invloed van incidentele absenties van de leerling. Met andere woorden, in toekomstig onderzoek behoeven uitvalcriteria gebaseerd op incidentele afwezigheid van de leerling, niet te worden toegepast op de verwerking van de scores op sociometrische tests;

2. De reden waarom de SAGS werd afgenomen, is dat groepscohesie als maat voor het samenwerkingsaspect uit de sociometrische structuur kan worden afgeleid.

test	group	class	pretest mean	pretest S.D.	posttest mean	posttest S.D.	n	diff.	pretest min.	pretest max.	posttest min.	posttest max.
SAGS O	exp.	2	2.55	.41	2.55	.47	18	.0	1.47	3.06	1.65	3.35
		3	2.15	.44	2.42	.67	16	+.27	1.53	2.87	1.33	3.87
	contr.	1	2.37	.35	1.84	.27	16	-.53	1.71	2.93	1.47	2.27
		4	1.59	.24	1.58	.26	21	-.01	1.30	2.10	1.10	2.15
SAGS G	exp.	2	2.55	.79	2.57	.83	18	-.02	1.00	3.76	1.00	3.74
		3	2.18	.56	2.42	.55	16	-.24	1.33	3.13	1.27	3.13
	contr.	1	2.37	.65	1.84	.65	16	+.53	1.29	3.43	1.00	3.73
		4	1.59	.42	1.59	.46	21	.0	1.05	2.40	1.05	2.65

Tabel 9.11. : Mean, standard deviation and range of the classes of the experimental and control group on the SAGS O and SAGS G, on pretest and posttest. Differences between post- and pretest.

De instructie van de SAGS is als volgt (Defares e.a., 1970):
'...Het kan wel eens gebeuren dat je je ongelukkig en verdrietig
'voelt. Bij voorbeeld als je heel erg teleurgesteld bent door iets,
'of als iemand erg onaardig tegen je geweest is. Denk je nu eens
'in dat je erg verdrietig was en dat je daar graag met iemand over
'zou willen praten. Nu moet je bedenken met wie je het liefst van
'alle mensen over je verdriet zou willen praten...'

Gezien de probleemstelling van het huidige onderzoek kan men zich afvragen of het wellicht relevanter is de instructie van de SAGS daarop af te stemmen.

A.B.V.

In tabel 9.12 zijn gemiddelde, standaarddeviatie, minimum en maximum van de ABV schalen vermeld, gemeten op de voortest en natest. Deze gegevens zijn overeenkomstig met die van het vooronderzoek.

scale	group	pretest mean	pretest S.D.	posttest mean	posttest S.D.	pretest min.	pretest max.	posttest min.	posttest max.	n
N	exp.	59.455	20.81	60.485	22.25	23	106	23	111	33
	contr.	63.486	17.54	67.351	21.22	32	99	23	107	37
NS	exp.	18.667	4.36	18.939	4.75	12	31	12	32	33
	contr.	19.946	5.75	21.459	4.85	12	33	14	29	37
E	exp.	56.818	11.64	57.455	12.70	32	86	33	81	33
	contr.	53.243	12.90	52.919	16.21	27	73	20	81	37
T	exp.	40.636	11.10	36.818	10.25	21	62	22	59	33
	contr.	38.703	9.19	36.216	7.68	24	60	21	56	37

Tabel 9.12. : Mean, standard deviation and range of the scores on the N-, NS-, E- and T-scale of the Amsterdamse Biografische Vragenlijst on pretest and posttest of the experimental and control group.

Omdat uit de factoranalyse (groep IV in 6.2.) bleek dat de factorstructuur van de schalen van de ABV in de experimentele en controlegroep overeenkomstig was (bovendien overeenkomstig aan de factorstructuur zoals Wilde (1970) vermeldt), werd op de N-, E- en T-scores een covariantieanalyse uitgevoerd (zie 6.3.).

Het bleek dat dit covariantieanalysemodel ten aanzien van de N- en de T-schaal een geringe verklaringskracht had; de beste schatting van het effect van de onafhankelijke variabele bleek voor de E-score een richting te hebben, tegengesteld aan het voorspelde teken (- in plaats van +). De niet aantoonbare veranderingen op deze persoonlijkheidskenmerken zijn in overeenstemming met de resultaten uit het vooronderzoek (Oud e.a., 1971). Dit kan hun interpretatie bevestigen dat blijkbaar de persoonlijkheidskenmerken die de ABV meet, te stabiel zijn om onder invloed van (extra) lessen lichamelijke oefening

ontstane veranderingen te detecteren.

In tegenstelling tot de gevonden resultaten in het onderzoek zijn de resultaten van Mulder (1962) te melden. Een groep proefpersonen tussen vijftien en twintig jaar nam gedurende vier weken deel aan een full-time trainingsprogramma op de Outward Boundschool te Renesse, waarin onder andere opdrachten moesten worden uitgevoerd die een beroep deden op kennis, intelligentie en karakter. Uit de scores op de schalen van de ABV bleek dat er significante veranderingen optraden op de N-, NS- en T-schaal (voortest-natest verschillen). Het verschil met de resultaten uit het onderhavige onderzoek zou kunnen worden verklaard door kenmerken van de trainingsprogramma's zelf, namelijk:

1. het expliciet gericht zijn op de verbetering van de prestaties in de lessen lichamelijke oefening in vergelijking met de intentie zoals Mulder die weergeeft: 'Men wil door middel van praktisch bezig zijn onmiddellijk laten beleven: gemeenschap, kameraadschap, hulpvaardigheid, initiatief, verantwoordelijkheid, zorgzaamheid, ordelijkheid'.
2. de spreiding van het programma over de tijd: vier weken full-time versus drie of vijf lessen lichamelijke oefening per week gedurende een schooljaar.

Gezien de relatief stabiele persoonlijkheidskenmerken die de ABV meet, kan men zich afvragen of de veranderingen welke Mulder vindt niet een 'short-term' trainingseffect zijn dat verdwijnen zal op langere termijn ('long-term' effect). Gezien de resultaten uit het vooronderzoek (Oud e.a., 1971) en de resultaten van het huidige onderzoek, zou men op grond van de doelstelling van het onderwijs in de lichamelijke oefening kunnen veronderstellen dat veranderingen die optreden in door middel van de ABV gemeten persoonlijkheidskenmerken te klein zijn om meetbaar gemaakt te worden. Indien deze veronderstelling juist is, zou longitudinaal onderzoek, waarbij de onafhankelijke variabele gedurende langere tijd aan de proefpersonen wordt aangeboden, uitsluitsel kunnen geven.

P.M.T. K.

In groep IV (zie 6.2.) waarin onder andere de PMT-K schalen P, F- en F+ zijn opgenomen, kon voor de experimentele en controlegroep geen overeenkomstige factorstructuur worden gevonden (zie 7.2.).

	experimental group (n=33)	control group (n = 37)	Hermans (n = 3887)
scale	P	P	P
F-	-.03	+.33**	-.24**
F+	-.19	-.00	+.23**
scale	F-	F-	F-
F+	-.48**	+.19*	-.61**

Tabel 9.13. : Intercorrelations of the PMT-K scales of the experimental and control group, and those found by Hermans (1971); (* = $p \leq .05$; ** = $p \leq .01$).

	experimental group (n =33)	control group (n = 37)	Hermans (n = 270)
scale	P	P	P
N	-.35**	-.38**	-.30**
NS	-.28**	-.36**	-.22**
E	+.23**	+.13	-.02
T	+.31**	+.30**	+.32**
scale	F-	F-	F-
N	-.01	+.38**	+.29**
NS	-.09	+.33**	+.33**
E	+.05	-.06	-.29**
T	+.17*	-.42**	+.08
scale	F+	F+	F+
N	+.03	-.19	-.30**
NS	+.14	-.05	-.32**
E	+.01	+.31**	+.18**
T	-.21*	+.13	-.10

Tabel 9.14. : Correlations of the PMT-K scales with the ABV scales of the experimental and control group, and those found by Hermans (1971); (* = $p \leq .05$; ** = $p \leq .01$).

Uit de intercorrelatiematrix in tabel 9.13. van de PMT-K schalen van de proefpersonen van de experimentele en controlegroep blijkt dat er aanzienlijke verschillen bestaan met de door Hermans (1971) gevonden intercorrelaties van de PMT-K schalen met die van de ABV tussen enerzijds de experimentele en controlegroep, en anderzijds tussen beide groepen met de door Hermans vermelde waarden (zie tabel 9.14.).

Gezien de afwijkingen met Hermans, en gezien de verschillen tussen de correlaties van de experimentele en controlegroep, lijkt het weinig zinvol om veel aandacht aan deze schalen van de PMT-K te besteden.

scale	group	pretest mean	pretest S.D.	posttest mean	posttest S.D.	pretest min.	pretest max.	posttest min.	posttest max.	n
P	exp.	18.576	4.56	15.727	6.34	8	28	4	29	33
	contr.	18.294	5.64	14.059	6.03	8	31	3	25	34
F-	exp.	8.212	3.00	8.091	3.60	1	14	2	15	33
	contr.	9.059	3.28	8.000	4.28	2	14	1	15	34
F+	exp.	9.788	3.79	9.594	4.29	3	17	2	17	32
	contr.	8.618	3.48	9.294	4.76	2	16	2	17	34

Tabel 9.15. : Mean, standard deviation and range of the scores of the P, F- and F+ scale of the PMT-K on the pretest and posttest of the experimental and control group.

Hoewel verschillen in correlaties kunnen voortvloeien uit het verschillend zijn van de experimentele en controlegroep, zijn hier echter geen indicaties voor te vinden (zie tabel 9.15.).

Twee algemene opmerkingen lijken echter gerechtvaardigd:

1. Hoewel een overdracht verondersteld wordt, maar niet aangetoond kan worden, kan men zich afvragen of het wellicht relevant is eerst na te gaan of in de les lichamelijke oefening zelf een verandering in de bereidheid tot het leveren van prestaties optreedt. Dit suggereert een specifieker instrument te zoeken;
2. De PMT-K is volgens Hermans vooral gericht op het meten van prestatiemotivatie op intellectueel gebied, met andere woorden, op

de prestatiemotivatie bij typische leervakken. Het lijkt daarom zinvol om pas dan de PMT-K als meetinstrument te hanteren wanneer ten aanzien van prestatiebereidheid in de les lichamelijke oefening zelf blijkt dat een verandering is opgetreden, en men op goede gronden zou mogen verwachten dat overdracht van deze verandering naar aspecten buiten de les lichamelijke oefening zal optreden.

S.S.D.

Ten aanzien van de tien concepten uit de SSD werd in de ongereduceerde multivariate analyse geen verschil gevonden tussen de experimentele en controlegroep. De tien concepten werden niet opgenomen in de covariantieanalyse, omdat in de experimentele en controlegroep geen overeenkomstige factorstructuur werd gevonden. Uit de voortestgegevens van de proefpersonen uit het vooronderzoek (Oud e.a., 1971) konden vijf factoren worden afgeleid (zie 4.2.). De twee hoogst ladende variabelen op iedere factor vormden de tien concepten van de SSD in dit onderzoek. Omdat het noodzakelijk was het aantal variabelen gezien het aantal proefpersonen te reduceren, werden de variabelen in groepen ingedeeld (6.2.). De tien concepten van de SSD werden beschouwd als één groep, waardoor wederom de mogelijkheid werd geschapen om een factoranalyse van deze variabelen uit te voeren. De veronderstelling dat de vijf factoren uit het vooronderzoek zouden verschijnen, lijkt aannemelijk. De veronderstelling werd echter niet bevestigd.

Drie opmerkingen voor dit verrassende doch teleurstellende resultaat kunnen geplaatst worden:

1. de ladingen op de factoren zijn zo laag dat geen eenduidige conslusies over een factorstructuur zijn te trekken, laat staan over een overeenkomstige factorstructuur;
2. de proefpersonen uit dit onderzoek kunnen verschillend zijn van de proefpersonen uit het vooronderzoek. De gegevens leveren geen aanwijzingen in die richting op;
3. het verschil in aantal adjectiefparen en de concepten kan van invloed zijn geweest.

Wat betreft het laatste punt kunnen de volgende kanttekeningen geplaatst worden. Uit incidentele observaties tijdens de afname van de test ontstond de indruk dat vooral bij de beoordeling van de eerste concepten de proefpersonen geen raad wisten met de scoring op de 7-puntsschaal bij sommige adjectiefparen. Dit zou kunnen wijzen op een niet-adequate instructie, maar de instructie was identiek aan die uit het vooronderzoek. Het wellicht belangrijkste verschil vinden we bij de te beoordelen concepten. In het vooronderzoek werden met de schalen uit de SSD geen tien maar twintig concepten beoordeeld, en bovendien werden deze concepten voorafgegaan door het concept: 'Ik zelf, zoals ik in werkelijkheid ben'. Dit concept werd gescoord op achtendertig andere adjectiefparen. Deze waren duidelijk toepasbaar op dit concept. Op grond van het voorgaande lijkt dit concept volgens ons in belangrijke mate te hebben bijgedragen tot een betere scoring van de twintig concepten op de SSD-schalen. Indien juist, dan zou dit kunnen betekenen dat in dit onderzoek de instructie op zich juist was, maar voor de jeugdige proefpersonen niet adequaat genoeg, met name onvolledig.

In aansluiting hieraan kan men zich bovendien afvragen of er niet een discrepantie bestaat tussen het voorbeeld in de instructie en de concepten welke ter beoordeling aan de proefpersonen worden voorgelegd (zie bijlage IV) zodat daardoor de scoring minder betrouwbaar wordt.

De semantische differentiaal is een flexibel instrument. Met andere woorden, men is vrij de adjectiefparen en de concepten te kiezen die men relevant acht voor de probleemstelling van het betreffende onderzoek. Men raadt echter aan om de concepten en ook de gekozen adjectiefparen te factoranalyseren om daarmee te controleren of de drie factoren van de semantische differentiaal, te weten: evaluatie, potentie en activiteit, wel in voldoende mate aanwezig zijn. Dit is tot dusver niet nagegaan.

Door drie psychologen en één leraar lichamelijke oefening werden in het vooronderzoek twintig concepten gekozen die volgens hen relevant waren met betrekking tot de doelstelling van de lichamelijke oefening, zoals geformuleerd in het Voorstel leerplan Rijksscholen.

De factoranalyse leverde weliswaar vijf factoren op, gerepresenteerd in vijf maal twee concepten met betrekking tot een aantal aspecten van deze doelstelling, maar in dit onderzoek werd dit resultaat niet gevonden. Blijkbaar is de gevolgde procedure geen garantie voor aanwezigheid van een stabiele factorstructuur.

Ten slotte wijzen we op de wenselijkheid dat de adjectiefparen in principe toepasbaar zijn op de te beoordelen concepten. Het gebruik van de SSD in de huidige vorm roept dienaangaande twijfels op. Voor proefpersonen, zoals in dit onderzoek, impliceert dit volgens ons dat objecten en subjecten te prefereren zijn boven abstracte begrippen.

Om een indruk te krijgen van de scores op de evaluatiefactor zijn deze per concept samengevat in tabel 9.16..

concept	group	pretest		posttest		pretest		posttest	
		mean	S.D.	mean	S.D.	min.	max.	min.	max
learning	exp.	22.970	3.95	20.303	4.50	12	28	10	28
	contr.	22.486	3.48	17.757	4.60	15	28	7	26
persever-ance	exp.	20.909	3.58	19.727	4.87	14	28	8	28
	contr.	22.270	3.54	19.946	4.42	14	28	13	28
cooper-ating	exp.	23.606	3.69	22.758	5.12	12	28	4	28
	contr.	23.541	3.04	20.432	4.78	13	28	4	28
getting up early	exp.	17.242	6.02	17.758	6.43	7	26	4	28
	contr.	18.108	6.34	16.344	7.41	4	27	4	28
hobbies	exp.	24.697	3.84	24.455	4.33	13	28	10	28
	contr.	24.154	2.99	22.541	5.61	16	28	7	28
perspiring	exp.	13.273	5.69	13.545	4.12	4	28	4	22
	contr.	12.892	4.89	13.054	4.26	4	22	4	19
sports	exp.	24.273	3.63	24.848	3.59	14	28	16	28
	contr.	25.090	3.76	22.432	6.50	13	28	4	28
helping each other	exp.	23.030	4.03	20.394	5.65	14	28	4	28
	contr.	23.351	3.52	18.892	5.44	14	28	7	28
arithmatics	exp.	21.606	5.41	18.394	6.96	8	28	4	28
	contr.	21.027	4.09	16.676	6.09	14	28	4	28
exertion	exp.	19.697	3.79	19.545	5.22	14	28	4	28
	contr.	19.811	4.11	18.234	4.81	10	28	10	28

Tabel 9.16. : Mean, standard deviation and range of raw pretest and posttest SSD-scores on the evaluation factor of the experimental and control group.

Schoolattitude

De covariantieanalyse toont ten aanzien van schoolattitude geen significante verandering bij de experimentele groep ten opzichte van de controlegroep. Op grond van de hypothese (7.3.) zou men een hogere score op schoolattitude verwachten onder invloed van lichamelijke oefening bij de experimentele groep ten opzichte van de controlegroep.

group	pretest		posttest		diff.	pretest		posttest		n
	mean	S.D.	mean	S.D.		min.	max.	min.	max.	
exp.	5.586	2.13	5.375	2.32	-.211	2	9	1	9	32
contr.	4.833	2.10	4.257	2.17	-.576	1	9	1	8	35

Tabel 9.17. : Mean, standard deviation, difference and range of pretest and posttest scores of the experimental and control group on school attitude.

Uit de gemiddelden van de voor- en natest (tabel 9.17) blijkt echter dat de score op de schoolattitudetest voor beide groepen zelfs afneemt. Deze daling is in overeenstemming met de resultaten van Flanders (1968) en Neal e.a. (1967). Zij vonden een daling ten aanzien van variabelen welke met school te maken hebben, zonder een manipulatie van variabelen welke relevant worden geacht voor een beïnvloeding van schoolattitude.

In dit onderzoek daarentegen bestaat de gemanipuleerde variabele uit het aantal lessen lichamelijke oefening. Het vak lichamelijke oefening wordt door de leerlingen als zeer positief ervaren (tabel 9.18.). Dit blijkt uit de hoge gemiddelde score op de schoolvoorkeurtest, namelijk een score van 6,9 bij een maximum van 9. De positieve waardering voor het vak lichamelijke oefening blijkt eveneens uit het gemiddelde rangnummer van het vak gymnastiek: 2,9, dus bijna op de derde plaats ten opzichte van de andere twaalf schoolvakken.

Omdat enerzijds een transfer van zo een positief gewaardeerd vak tot een toename van schoolattitude zou kunnen leiden, en anderzijds op grond van voornoemde literatuur een daling van

schoolattitude verondersteld mag worden, zou men kunnen veronderstellen dat lessen lichamelijke oefening de daling in schoolattitude zou kunnen compenseren. Het verschil in daling tussen de experimentele en de controlegroep in dit onderzoek zou dan eventueel verklaard kunnen worden door de onafhankelijke variabele.

Een tweede opmerking over het gevonden resultaat betreft het feit dat de proefpersonen in dit onderzoek brugklassers zijn. Met andere woorden: van deze leerlingen zou men kunnen aannemen dat zij met een hoge verwachting van de lagere naar de middelbare school gaan. Tijdens het schooljaar doen zij een dosis schoolervaring op (Kuethe, 1968), waardoor een aanpassing aan de gestelde verwachtingen optreedt, en wel zodanig dat de eventuele positieve transfereffecten van de lichamelijke oefening worden gecompenseerd.

Indien de veronderstelling juist is dat de schoolattitude wordt beïnvloed door elkaar tegenwerkende factoren, kan men zich afvragen of het niet zinvoller is een schoolattitudetest te ontwikkelen waarin deze twee factoren als afzonderlijke variabelen kunnen worden gescoord.

Gymnastiek voorkeur en gymnastiek rangnummer

test	group	pretest		posttest		diff.	pretest		posttest		n
		mean	S.D.	mean	S.D.		min.	max.	min.	max.	
preference of physical education	exp.	6.909	1.44	7.061	1.41	+.152	3	8	3	8	33
	contr.	6.919	1.36	7.135	1.21	+.236	2	8	4	8	37
rank of physical education	exp.	2.939	2.54	2.545	2.67	-.394	1	11	1	12	33
	contr.	2.703	2.90	2.459	2.22	-.244	1	13	1	8	37

Tabel 9.18. : Mean, standard deviation, and range of scores on test for preference and test for ranking of physical education at school, on pretest and posttest of experimental group and control group. Differences between post- and pretest.

Omdat uit de scores in tabel 9.18. blijkt dat de lichamelijke oefening reeds op de voortest een zeer geliefd vak is, is het moeilijk om een

verbetering aantoonbaar te maken. Bovendien zijn de scores afhankelijk van de andere schoolvakken, waarbij dezelfde vakken door verschillende leraren gegeven worden, en met verschillende onderwijsmethoden. Het is daarom waard te overwegen om ieder vak apart door de proefpersonen te laten scoren op een n-puntsschaal, waarvan de polen zijn afgestemd op de doelstelling van het betreffende onderzoek. Dit zou een suggestie kunnen zijn om het tevens mogelijk te maken de transferveronderstelling, namelijk de vraag of overdracht van het effect van een gemanipuleerde variabele naar andere schoolvakken plaatsvindt, meetbaar te maken.

9.7. BIOLOGISCHE LEEFTIJD

De biologische leeftijd van de proefpersonen werd zowel op de voortest als op de natest gemeten.

Zoals in de resultaten van hoofdstuk 5.1. reeds werd vermeld, bleek de toename in skeletleeftijd geen enkele correlatie te vertonen, niet met de skeletleeftijd op de voortest, noch met de kalenderleeftijd op de voortest.

Om de biologische leeftijd in de covariantieanalyse te kunnen opnemen, moest tevoren een keuze worden gemaakt uit de volgende alternatieven: skeletleeftijd op de voortest, skeletleeftijd op de natest, of de toename van de skeletleeftijd tussen voor- en natest.

Hoewel bekend is dat de biologische leeftijd niet wordt beïnvloed door lichamelijke activiteit zelf (Cerny, 1969), hebben wij besloten om uitsluitend de skeletleeftijd op de voortest te gebruiken als karakteristiek van de storende variabele : biologische leeftijd.

Alle skeletfoto's werden na een voorafgaande trainingsperiode door dezelfde persoon beoordeeld; uit de ruwe score werd vervolgens de leeftijd vastgesteld. Alle foto's van zowel de voor- als natest werden tegelijk en door elkaar heen beoordeeld. In enkele gevallen deed zich de moeilijkheid voor dat bij een proefpersoon de score van de natest lager was dan van de voortest. In deze gevallen hebben wij de 'toename' in skeletleeftijd op nul maanden gesteld. Een dergelijke moeilijkheid kan grotendeels vermeden worden door

niet de skeletleeftijd, maar de ruwe score als score te nemen. Bij bepaalde skeletleeftijden treedt een relatief grote verandering op in skeletleeftijd, bij een geringe verhoging van de ruwe score.

9.8. BUITENSCHOOLSE LICHAMELIJKE ACTIVITEIT

De resultaten van de metingen van de buitenschoolse lichamelijke activiteit geven grote inter-individuele verschillen te zien (zie tabel 5.10.). Dit heeft ongetwijfeld consequenties voor de invloed van de onafhankelijke variabele. Om hiervan een indruk te krijgen, kan men de stappenscores van de lessen lichamelijke oefening (tabel 5.21.) vergelijken met die van de vrije tijd.

De gemiddelde stappenscore tijdens de lessen bedraagt ongeveer 3.000, wat neerkomt op een gemiddelde totaalscore van 6.000 voor twee extra lessen lichamelijke oefening per week. De geschatte wekelijkse buitenschoolse activiteit van een vrij inactieve leerling is uitgedrukt in stappenscores 24.000 plus 9.000 geschat voor de drie lessen lichamelijke oefening per week die in het normale rooster zijn opgenomen. Dit betekent dat de twee extra lessen, uitgedrukt in stappenscores, een verhoging van de totale activiteit met ongeveer 18% geven. Bij zeer actieve leerlingen kan de totale score op de stappentellers over een hele week geschat worden op 180.000 plus 9.000. Dezelfde twee extra lessen betekenen in dat geval een verhoging van de totale activiteit per week met slechts 3 %.

Het lijkt aannemelijk dat in voornoemde gevallen de invloed van twee extra lessen op de ontwikkeling zal verschillen. Voor deze leeftijdfase, waarin de kinderen nog regelmatig buiten spelen en vrijwel allemaal naar school fietsen, zijn twee lessen meer per week voor de meer actieve leerlingen waarschijnlijk onvoldoende om veranderingen te veroorzaken. Het is echter de vraag of ook bij oudere leerlingen het gemiddelde activiteitsniveau nog zo hoog zal liggen en de onderlinge verschillen zo groot zullen zijn. In dat geval zouden twee lessen per week extra misschien wel effect hebben.

De relatie tussen de scores op de twee instrumenten die werden gebruikt voor het meten van de lichamelijke activiteit van de leerlingen in de vrije tijd, te weten, de vragenlijst afgenomen door een interviewer, en de stappentellers, was 0,50 ($p \leq 0,01$). Het is echter niet verwonderlijk dat deze relatie niet beter is, gezien het feit dat beide instrumenten niet hetzelfde hebben gemeten. Dit komt voort uit de beperkingen in de registratie met behulp van de stappentellers. Bevestigd aan de broekband reageren deze namelijk niet op alle vormen van lichaamsbeweging, maar alleen op die bewegingen waarbij zwaartepuntverplaatsingen in verticale richting optreden. Dit heeft tot gevolg dat met name het fietsen, een activiteit die door vrijwel alle leerlingen tamelijk veelvuldig werd beoefend en dan vooral in verband met het transport van en naar school, niet door stappentellers wordt geregistreerd. Aangezien uit de vragenlijst blijkt dat dit transport gemiddeld 30,8 % van de totale lichamelijke activiteit van de leerlingen uitmaakte, kan van een volledige vergelijking van de scores al geen sprake zijn. Juist in verband hiermee vormt de vragenlijst dan ook een belangrijke aanvulling op de meting van de lichamelijke activiteit met behulp van stappentellers.

Aangezien bouw en samenstelling van het lichaam kunnen veranderen onder invloed van lichamelijke inspanning, en deze eventuele veranderingen afhankelijk zijn van de mate van activiteit, verwachtten wij een relatie te vinden tussen de scores voor de buitenschoolse lichamelijke activiteit en de diverse antropometrische maten (tabel 9.19). Uit deze tabel blijkt dat op de voortest deze relatie over het algemeen bestaat, waarbij de correlatiecoëfficiënten niet erg hoog zijn. De correlaties met vet, spier, gewicht en breedtematen zijn over het algemeen negatief, hetgeen erop zou duiden dat leerlingen die in de vrije tijd actief zijn, smal gebouwd en mager zijn, ofwel het type van de duuratleet vertonen.

Voor het verkrijgen van trainingseffecten is behalve de totale hoeveelheid ook de intensiteit van de activiteit van groot belang. Aangezien de stappentellers en vragenlijst echter geen informatie verstrekken over de intensiteit van de buitenschoolse activiteiten, is in het geval van identieke scores niet duidelijk of deze tot

Correlations between habitual physical activity and anthropometric variables		
measurement	pedometer	questionnaire
weight	-.40**	-.05
height	-.42**	-.02
biacromial diameter	-.35**	-.06
biiliocristal diameter	-.26*	-.05
upper arm circumference	-.32**	-.13
thigh circumference	-.29**	-.10
calf circumference	-.40**	-.14
wrist breadth	-.24*	+.05
bicondylar femur	-.28**	+.04
bicipital skinfold	-.29**	-.17
tricipital skinfold	-.17	-.08
subscapular skinfold	-.27*	-.09
suprailiacal skinfold	-.21*	+.05
corrected upper arm diameter	-.29**	-.13
% fat	-.20*	-.07

Tabel 9.19 : Spearman correlation coefficients between scores of habitual physical activity and pretest scores of all anthropometric variables (n = 70); (* = $p \leq .05$ ** = $p \leq .01$).

stand zijn gekomen door activiteiten met een hoge intensiteit gedurende een korte periode, of door activiteiten met een veel lagere intensiteit, maar over een veel langere tijd volgehouden (De Boorder, 1971). De aanpassing van het lichaam zal bij deze identieke scores afhankelijk zijn van de intensiteit.

De relatie tussen activiteit in de vrije tijd en een aantal lichaamsmaten zou vermoedelijk beter zijn indien de intensiteit van de activiteiten in de scores tot uitdrukking zou komen.

Voortredenerend in dezelfde lijn werd verondersteld dat de mate van buitenschoolse lichamelijke activiteit in verband zou staan

met het lichamelijk prestatievermogen. De prestaties op de gymnologische tests bleken echter niet met de buitenschoolse activiteit te correleren.

Van de eenvoudige prestatietests was alleen de correlatiecoëfficiënt van de stappenscore met 'bent arm hang' zowel bij de experimentele als de controlegroep significant (tabel 9.20.).

Correlations between habitual physical activity and physical performances		
measurements	pedometer	questionnaire
plate tapping	-.13	-.08
sit and reach	+.27*	+.10
50 m. shuttle run	-.08	-.12
bent arm hang	+.34**	+.23*
vertical jump	+.07	+.02
handgrip	-.11	+.08
12 min. run walk	+.31**	+.30**
W_{170}	+.21*	+.35**
FEV %	-.11	-.17

Tabel 9.20.: Spearman correlation coefficients of scores of habitual physical activity with pretest scores of physical performance tests (n = 70);
(* = $p \leq .05$ ** = $p \leq .01$).

Dit zou erop kunnen wijzen dat de lichamelijke aanleg van kinderen van meer betekenis is dan de mate van lichamelijke activiteit, en dat meting van die activiteit gedurende een zo korte periode van het leven, zonder daarbij tevens de intensiteit vast te stellen, onvoldoende informatie verstrekt.

Het rookgedrag van de leerlingen bleek, wanneer de vragenlijst als score werd gehanteerd, significant te correleren met de buitenschoolse activiteit. De correlatiecoëfficiënt was echter positief, hetgeen betekent dat juist de actieve leerlingen meer rookten.

Gezien het geringe percentage van de leerlingen die meedeelden dat ze rookten, mag aan deze gegevens echter niet teveel waarde worden gehecht.

9.9. BEOORDELING LESGEEFGEDRAG

Binnen de experimentele groep werd lesgegeven door de leraren, respectievelijk B en C; binnen de controlegroep door de leraren A en D. Om het verschil in lesgeefgedrag van leraren te registreren en te kwantificeren, werd dit beoordeeld op vaktechnische en sociaal-emotionele aspecten, en ook door middel van een vragenlijst die werd voorgelegd aan de leerlingen. Bovendien werd als een zeer belangrijk aspect van het lesgeven van elke leraar de arbeidsintensiteit geschat met behulp van stappentellers. De resultaten van deze metingen wijzen uit dat me de laatstgenoemde methode geen significante verschillen tussen de klassen konden worden aangetoond (zie 5.3.3.).

De eventuele invloed van het leraargedrag op leerlingvariabelen is verder afgezwakt doordat de zwemlessen, welke een derde deel van het totale lesprogramma lichamelijke oefening in beslag namen, niet in klasverband werden gegeven; de leerlingen waren echter in prestatiegroepen ingedeeld zodat zij niet steeds van de eigen klasleraar les kregen.

Bij het ontwerpen van de interactieanalysesysteem is men ervan uitgegaan dat dit instrument geschikt moest zijn voor meer algemeen gebruik en niet alleen voor de vraagstelling in dit onderzoek. Zo zou men bij voorbeeld uit het systeem de activiteit van de leerlingen (collectief) kunnen afleiden, of de inhoud van de instructie van de leraar nader kunnen analyseren. In dit onderzoek werden hiervoor nauwkeuriger en adequater instrumenten gebruikt, zoals de stappentellers en de vaktechnische beoordeling.

Vaktechnische beoordeling

In feite was het vaktechnische beoordelingssysteem gekozen om een zo gedifferentieerd mogelijke indruk te krijgen over hoe men in de vaktechnische wereld het lesgeefgedrag kan beoordelen. De tien

subcategorieën waren met opzet zo gekozen dat de formulering aansloot bij het taalgebruik in de vakwereld, wat tevens inhield dat de omschrijvingen van de categorieën ruim geïnterpreteerd konden worden. Dit kon men onder meer afleiden uit de resultaten verkregen bij de bepaling van de reproduceerbaarheid. Het bleek dat de beoordelaars de tweede beoordeling lager (strenger) scoorden dan de eerste beoordeling van diezelfde les (de verschillen waren soms zeer groot: maximaal 3 punten op een 5-puntsschaal), terwijl de beoordelaars zelf verklaarden dat ze het idee hadden de tweede beoordeling hoger (milder) te hebben gescoord. Dit wekte bij de onderzoekers enige verbazing, maar een duidelijke verklaring kon men niet geven (zie 6.1.).

De rangvolgordes van de leraren waren voor alle categorieën vrijwel gelijk (zie 5.3.1.). Dit zou betekenen dat de afzonderlijke leraren op alle categorieën even 'kundig' of 'onkundig' zouden zijn, hetgeen niet aannemelijk is. Dit is mogelijk te verklaren doordat de categorieën van dit beoordelingssysteem een 'high inference' karakter hebben en daardoor gevoelig zijn voor vooroordelen, bij voorbeeld in de vorm van 'halo effecten'. Dit wil zeggen dat er ruime interpretatiemogelijkheden voor de beoordelaar bestaan en dat alle scores die een beoordelaar geeft, subjectief worden beïnvloed door een algemene evaluatie van de leraar wanneer deze in het begin van een lesonderdeel een goede of een slechte indruk maakt (Krech e.a., 1969). Deze veronderstelling wordt bevestigd door het feit dat de beoordelaars het niet nodig vonden om een lesonderdeel in zijn geheel uit te zien om de scores te bepalen. Tevens kon tijdens de beoordeling niet altijd contact tussen de beoordelaars worden voorkomen, waardoor uitroepen van goed- of afkeuring de andere beoordelaars kunnen hebben beïnvloed. Bij een van de leraren kan nog een rol hebben gespeeld dat hij mogelijk als de projectleider van het onderzoek is herkend. Ter voorkoming hiervan is het contact met de beoordelaars zoveel mogelijk door andere medewerkers aan het onderzoek onderhouden.

Concluderend mogen we opmerken dat het aanbeveling verdient om ernaar te streven dat de vaktechnische aspecten van het lesgeefgedrag ook met een 'low inference' categorieënsysteem worden gescoord.

Sociaalemotionele beoordeling

In het onderzoek naar het sociaalemotionele gedrag van de leraar tijdens de lessen lichamelijke oefening waren twee fasen van belang. Allereerst moest een instrument worden ontwikkeld om het klimaat in de klas te bepalen, omdat dit instrument in Nederland voor de lessen lichamelijke oefening nog niet bestond. Daarna kon pas met het werkelijke onderzoek naar het leraargedrag worden begonnen.

Na de start van het totale project werd tot dit extra onderzoek besloten, in de wetenschap dat door tijdgebrek problemen zouden optreden bij:

1. het normeren, valideren en op betrouwbaarheid onderzoeken van het instrument;
2. het trainen van beoordelaars;
3. het doen van literatuuronderzoek vooraf.

ad 1. Dat niet genormeerd is, heeft tot gevolg dat verschillen tussen de leraren op het continuum directief-nondirectief geen informatie geven over de absolute positie van de leraren ten opzichte van elkaar, alsmede ten opzichte van extremen op bovengenoemd continuum. De effecten van de geconstateerde verschillen kunnen niet, of moeilijk, in verband gebracht worden met leerlinggedrag omdat

a. vergelijking met ander onderzoek, waarin zowel verschillen alsook effecten daarvan op het leerlinggedrag werden onderzocht, biedt geen oplossing omdat daarin meestal wordt gewerkt met extreem scorende leraren op bovengenoemd continuum;
b. bovendien werd slechts ander onderzoek gedaan betreffende de zogenaamde 'zaakvakken'. Daarin komt men tot de conclusie dat meer non-directief lesgeefgedrag gunstiger inwerkt op leerlingvariabelen. Voor de lessen lichamelijke oefening hoeft deze tendens niet te gelden omdat in deze lessen door de meeste leraren veel instructie wordt gegeven en een groot deel van de les wordt besteed aan het zelf oefenen (het 'huiswerk' bij de zaakvakken). Tevens kan de grote ruimte waarin wordt lesgegeven (in de zaal of in de buitenlucht) een rol spelen; ook is de aard van de oefenstof geheel verschillend. Het directieve of non-directieve klimaat in de les lichamelijke oefening heeft hierdoor waarschijnlijk een andere

betekenis dan in de andere lessen.

In dit onderzoek bleken de beoordeelde lessen voor 80% uit instructie te bestaan (categorie 7, 9, 10 en 14). De score op de resterende categorieën, van belang voor het bepalen van het sociaal-emotionele klimaat, namen 20% van de tijd in beslag (zie tabel 5.20.). Op deze categorieën namen de leraren B en C uit de experimentele groep ten opzichte van elkaar een extreme positie in. De leraren A en D bevonden zich hier meestal tussenin, maar dichter bij leraar C dan bij leraar B. Bij de vaktechnische en leerlingbeoordeling is hetzelfde beeld terug te vinden: leraar B en C nemen een extreme positie in, leraar A en D bevinden zich echter meer in de buurt van leraar B (zie tabel 5.13.). Uit de resultaten van het onderzoek naar eventuele klas- en/of leraareffecten (9.1.) blijkt dat deze op een aantal afhankelijke variabelen waarschijnlijk van invloed zijn geweest. Waneer effecten aan de leraar mogen worden toegeschreven, zouden deze veroorzaakt kunnen zijn door de grote verschillen tussen de leraren B en C. In dat geval bestaat de mogelijkheid dat bij de experimentele groep het netto effect van de onafhankelijke variabele op de afhankelijke variabele mogelijk is verminderd door 'tegengestelde' effecten, veroorzaakt door de extreme posities van de leraren die lesgaven aan deze groep.

ad 2. Door tijdgebrek kon de reproduceerbaarheid van de beoordelaars niet meer worden bepaald vóór de aanvang van de observatieperiode; dit is alsnog gedaan bij drie van de vijf beoordelaars na de beoordeling van alle lessen. Het gemiddelde van de verkregen reproduceerbaarheidscoëfficiënten was 0,80, met een minimum van 0,67 en een maximum van 0,91 (zie tabel 6.4.).

Om een eventuele te lage objectiviteit te compenseren, heeft men vanaf het begin gewerkt met vijf beoordelaars. Zodoende konden per categorie en per beoordelaar extreme scoringen worden geëlimineerd. Om evenwel te grote onderlinge afwijkingen te voorkomen, hebben de beoordelaars in het begin van de observatieperiode overleg gepleegd over de onderlinge verschillen in het scoren op sommige categorieën. Hierbij werd gebruikgemaakt van een aantal korte verslagen, die overigens na iedere beoordeling werden opgesteld en waarin storingen werden vermeld en opmerkingen over de les werden gemaakt. Omdat

de verschillen van systematische aard leken, werd besloten dat de beoordelaars tussentijds niets zouden veranderen aan de manier waarop zij de gedragscategorieën scoorden. In hoeverre na middeling van de uitkomsten toch nog in de resultaten een vertekend beeld van de werkelijkheid bleef bestaan, is niet na te gaan.

Wanneer in volgend onderzoek een goede training vooraf mogelijk is waardoor bij de beoordelaars een hoge objectiviteit en reproduceerbaarheid ontstaat, kan worden volstaan met een of twee beoordelaars per les. Dit betekent een grote tijdwinst wanneer veel lessen moeten worden beoordeeld (Medley e.a., 1958).

ad 3. Uit ander onderzoek blijkt dat in het door ons ontwikkelde interactieanalysesysteem verbeteringen kunnen worden aangebracht. De volgende suggesties zijn van toepassing:

In dit systeem wordt verbaal en niet-verbaal gedrag gescoord van zowel leraar als leerlingen. Dit geeft problemen omdat twee soorten gedrag kunnen samenvallen in tijd en ruimte. Het Flanderssysteem bij voorbeeld (Flanders, 1970) is een geheel verbaal systeem, en als men ervan uitgaat dat per tijdseenheid slechts één persoon spreekt, kunnen de categorieën elkaar niet overlappen. Om overlapping in ons systeem te voorkomen, werd primair het leraargedrag gescoord, met als gevolg dat leerlingcategorieën de werkelijkheid onvolledig weergeven. Een oplossing zou zijn: a. twee beoordelaars aanwijzen waarvan één het leraargedrag beoordeelt, en één het leerlinggedrag; b. één beoordelaar die korte periodes in dezelfde les afwisselend op beide systemen scoort (Medley e.a., 1963).

Het aantal categorieën in een systeem mag niet te groot zijn, omdat dan de beoordelaars niet meer snel en accuraat kunnen scoren. De meeste onderzoekers gebruiken niet meer dan tien categorieën, terwijl in dit interactieanalysesysteem zeventien categorieën worden gehanteerd. Uit ander onderzoek blijkt dat bij gebruik van meer categorieën toch snel en accuraat gescoord kan worden wanneer een eenvoudig scoringssysteem wordt gebruikt: bij dit onderzoek diende slechts één toets van de 'teletype' te worden ingedrukt om een categorie voor langere tijd te kunnen scoren.

Tevens blijft het bij een zeventien-categorieënsysteem mogelijk

om later enkele categorieën samen te voegen indien ter ontwikkeling van een systeem mocht blijken dat sommige categorieën statistisch niet hanteerbaar zijn. Wanneer dit wordt gedaan, moet echter rekening worden gehouden met twee dingen:

a. het is wenselijk om op grond van statistische overwegingen de categorieën zodanig te definiëren dat hun frequenties zoveel mogelijk over alle categorieën gelijkelijk zijn verdeeld. In dit systeem zou dit betekenen dat elke categorie ongeveer 6% van het totale lesgebeuren in beslag moet nemen. In het gehanteerde systeem bleek dit niet het geval. Medley e.a. (1963) verklaren echter dat categorieën - ongeacht in welk systeem - die minder dan 5% van het totale gedrag in beslag nemen, ook nog goed kunnen functioneren;

b. de categorieën in dit systeem zijn enerzijds voor een deel ontstaan door factoranalyse die door andere onderzoekers is uitgevoerd op het leraargedrag, anderzijds zijn de categorieën voor een deel ontworpen op grond van theoretische overwegingen. De laatste op grond van theorie ontworpen categorieën hoeven niet statistisch onafhankelijk te zijn en kunnen elkaar daardoor overlappen. Dit kan problemen opleveren bij de statistische verwerking van de gegevens, wanneer onafhankelijkheid wordt vooronderstelt; in een systeem met veel categorieën kan dit worden ondervangen door afhankelijke categorieën samen te voegen.

Om wederzijdse beïnvloeding der categorieën te voorkomen, waardoor betere statistische bewerkingen van de gegevens mogelijk worden (Medley e.a., 1963), verdient het voorkeur dat de gebeurtenissen die voor het systeem van belang zijn, worden geteld en niet - zoals hier - worden gescoord als tijdsduur. Men telt dan het aantal malen op dat een leraar bij voorbeeld instructie geeft, onafhankelijk van hoe lang dit telkens duurt. Men kan uiteindelijk, indien nodig, gewichten toekennen aan de verschillende categorieën. In het gehanteerde systeem kan uit de frequentiecellen van de matrix indirect het aantal malen geteld worden dat een categorie plaatsvond.

Door verscheidene onderzoekers is opgemerkt (Tausch e.a., 1967; Wallen e.a., 1963) dat zowel goed- als afkeuring, gegeven door de leraar, irreversibele gedragingen zijn en afhankelijkheid van de leerling ten aanzien van de leraar veroorzaken. Tevens leidt individuele

lof tot verminderde sociale integratie van de leerlingen, doordat zij hun aandacht meer richten op de leraar. Bovendien hebben goedkeuring en afkeuring psychologisch gezien een reinforcement karakter met betrekking tot gedragscontrole (Becker e.a., 1967). Deze controle zou beperkend werken op de vrijheid van de leerling ten aanzien van initiatief en creativiteit.

In het gehanteerde systeem wordt individuele goedkeuring geïnterpreteerd als non-directief gedrag, dat wil zeggen dat de leraar reversibel met de leerlingen omgaat en dat aan de leerlingen veel ruimte wordt geboden tot initiatief in de vorm van het opperen van eigen ideeën etc.. Dit is in tegenspraak met het bovenstaande, en er moet dus nogmaals worden onderzocht onder welke gedragseenheid bepaalde vormen van goedkeurend gedrag moeten worden ingedeeld.

Leerlingbeoordeling

De betrouwbaarheid van de vragenlijst is door tijdgebrek niet nagegaan, waardoor de resultaten ervan met enig voorbehoud moeten worden geïnterpreteerd. Uit de resultaten blijkt dat de leerlingen in hun oordeel differentiëren tussen persoonlijkheid en vakmanschap van de leraar: de leerlingen beoordeelden leraar B het laagst op het eerstgenoemde aspect (schaal 1, 2 en 3), maar het hoogst op het tweede aspect (schaal 4 en 6). Dit is in tegenspraak met het onderzoek van Cogan (1963) die bij het afnemen van vragenlijsten aan veertienjarige leerlingen op verschillende middelbare scholen een hoge samenhang vond (met behulp van covariantieanalyse) tussen 'inclusive' en 'competence' gedrag van de leraar. Beide zijn gedragsaspecten die overeenkomen met respectievelijk schaal 1, 2 en 3, en schaal 4 en 6. Een mogelijke verklaring van deze verschillen is misschien dat de steekproef van Cogan uit 33 leraren bestond (987 leerlingbeoordelingen). In tabel 9.21. worden de resultaten van de leerlingbeoordeling vergeleken met min of meer overeenstemmende categorieën uit het interactieanalysesysteem en het vaktechnische beoordelingssysteem.

Het blijkt dat de leraren B en C op alle hiergenoemde kenmerken extreme posities ten opzichte van elkaar innemen, behalve op schaal

System of judgement	Category(ies) or scale(s)	Teacher							
		A		B		C		D	
		rv	mean	rv	mean	rv	mean	rv	mean
professional skills	total of all 10 categories	3	3.32	2	3.42	4	2.84	1	3.51
	category 3.2: joy	1.5	3.60	1.5	3.60	4	3.17	3	3.58
social-emotional aspects	ratio: acceptation (cat. 1+2) criticism (cat. 12+13)	2	2.6	1	6.9	4	0.95	3	2.4
	pupil initiation (cat.6)	2	371	4	115	1	472	3	246
judgement by pupils	scale 1+2+3 (soc.emot. aspects)	2	3.6	4	2.7	1	4.9	3	3.0
	scale 4+6 (prof.skills)	4	4.1	1	5.9	2	5.7	3	4.8

Tabel 9.21. : Comparison of results from judgements of teaching behaviour of three different systems: judgement of professional skills, of social-emotional aspects, and judgements by pupils;
rv = rank order between teachers;
mean = mean scores of subsequent judgements.

4 en 6 van de leerlingbeoordeling, waarin beide leraren als hoogste scoren. De vaktechnische beoordelaars en de leerlingen achten leraar B vaktechnisch het meest bekwaam. Ook heeft hij de hoogste waarde op de ratio: prijzen/critiek. Hij geeft echter de leerlingen het minst de mogelijkheid tot het nemen van initiatief, en de leerlingen mogen hem als persoon het minst, terwijl zij, blijkbaar terecht, erkennen dat hun leraar goed lesgeeft. Dit in tegenstelling tot leraar C. Deze wordt het minst bekwaam geacht door de vaktechnische

beoordelaars, geeft de meeste critiek en wordt toch door de leerlingen beschouwd als de aardigste, en als bijna even kundig als leraar B. De 'rommelige' lessfeer, genoemd in de beoordelingverslagen van de beoordelaars van het sociaalemotionele klimaat, en de mogelijkheden tot het tonen van initiatief, zijn blijkbaar (voor een deel) bepalend voor het oordeel van de leerlingen. Dit komt overeen met de resultaten uit het onderzoek van Veen (1969) die met enige verbazing constateerde dat in een rommelige en rumoerige groep de leraar door de leerlingen in sociaalemotioneel opzicht positief werd beoordeeld, terwijl de observatoren (werkstudenten) dit niet hadden verwacht. In ons onderzoek lieten de beoordelaars van het sociaalemotioneel beoordelingssysteem in hun verslagen ook blijken dat leraar C onoverzichtelijk en het minst plezierig lesgaf, terwijl de vaktechnische beoordelaars vonden dat de leerlingen, waarschijnlijk op grond van dit lesgeefgedrag, de minste arbeidsvreugde zouden ondervinden. Voor de leerlingen waren deze aspecten van het lesgeefgedrag blijkbaar niet maatgevend, of er nu veel of weinig goedkeurende opmerkingen werden gedaan. Dit kan misschien in verband gebracht worden met hetgeen Tausch e.a. (1967) en Walleen e.a. (1963) hebben gezegd over 'goedkeuring' van de leraar, namelijk, dat dit afhankelijk gedrag bij de leerling veroorzaakt. De leerlingen ondergaan dit waarschijnlijk als minder plezierig en vinden het fijner om in een minder strak geleide situatie les te krijgen.

Leraar B en leraar C nemen op schaal 4 en 6 geen extreme positie ten opzichte van elkaar in en hebben de hoogste gemiddelde score ten opzichte van de leraren A en D. Leraar B en leraar C gaven les aan de experimentele groep. De vragenlijst van de leerlingbeoordeling over deze leraren werd afgenomen op de natest. Een mogelijke verklaring voor deze resultaten ligt in de veronderstelling dat inhoud en frequentie van de lessen van invloed is geweest op het oordeel van de leerlingen. In de ogen van de leerlingen konden de leraren uit de experimentele groep meer 'leuke dingen' doen dan de leraren uit de controlegroep, waardoor zij hun eigen leraar hoger beoordeelden.

Literatuur

1. Becker, W.C., C.H. Madsen, C.R. Arnold, D.R. Thomas, The contingent use of teacher attention and praise in reducing classroom behavior problems, J. Special Educ., 1, 3, 1967, 287-307.
2. Boorder, Tj. de, Energiebesteding en lichamelijk prestatievermogen, Proefschrift Universiteit van Amsterdam, Elkerbout, Scheveningen, 1971.
3. Cogan, M.L., The behavior of teachers and the productive behavior of their pupils, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1963, 1031-1059.
4. Cumming, G.R., Correlation of athletic performance and aerobic power in 12 to 17 year old children with bone age, calf muscle, total body potassium, heart volume and two indices of anaerobic power, in: Pediatric Work Physiology, Proc. 4th. Int. Symp., ed.: O. Bar-Or, Natanya, 1973, 109-134.
5. Defares, P.B., G.N. Kema, E. van Praag, J.J. van der Werff, Syracuse-Amsterdam-Groningen Sociometrische Schaal, uitgave voor researchdoeleinden, 1970.
6. Flanders, N.A., B.M. Morrison, E. Leland Brode, Changes in pupil attitudes during the school year, Journal of Ed. Psychology, 50, 1968, 334-338.
7. Flanders, N.A., Analysing teaching behavior, Addison-Wesley publishing company, Massachusetts, 1970.
8. Hermans, H.J.M., Prestatiemotivatie en Faalangst in Gezin en Onderwijs, Tevens handleiding bij de Prestatie Motivatie Test voor Kinderen (P.M.T.-K), Swetz en Zeitlinger, Amsterdam, 1971.
9. Kemper, H.C.G., Heart rate during bicycle ergometer exercise in Watts per kilogram body weight of 12 and 13 year old boys, Z. Kinderheilk. 113, 1972, 161-169, (a).
10. Kemper, H.C.G., Experimenten over de lichamelijke aspecten van de lichamelijke opvoeding bij 12 en 13-jarige jongens, Richting, 12, 1972, 236-243, (b).
11. Klissouras, V., G. Weber, Training, growth and heredity, in: Pediatric Work Physiology, Proc. 4th Int. Symp., ed: O. Bar-Or, Natanya, 1973, 209-215.

12. Krech, D., R.S. Crutchfield, N. Livson, Elements of psychology, Alfred A. Knopf Inc., New York, 1969.
13. Kuethe, J.L., The teaching learning process, Glenview III, Scott, Foresman and Co., 1968, 92.
14. Medley, D.M., H.E. Mitzel, Application of analysis of variance to the estimation of the reliability of observations of teachers' classroom behavior, J. Exp. Educ., 27, 1958, 23-25.
15. Medley, D.M., H.E. Mitzel, Measuring classroom behavior by systematic observation, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1963, 247-328.
16. Mulder, H.F., Karaktervorming door Lichamelijke Inspanning, Peadagogische studiën, 39, 12, 1962.
17. Neal, D.C., J.M. Proshek, School related attitudes of culturally disadvantaged school children, Journal of Ed. Psychology, 1967, 238-244.
18. Oud, A.N.P., L.W.C. Tavecchio, Een onderzoek betreffende de invloed van extra lessen lichamelijke opvoeding op leerlingen uit de brugklassen van een school voor VWO en HAVO op lichamelijke en geestelijke groei en ontwikkeling, Werkstuk, Laboratorium voor Psychofysiologie, Universiteit van Amsterdam, 1971, en in: Kemper, H.C.G. e.a., De Lich. Opv., 59, 12 en 13, 1971, 543 en 604.
19. Tausch, R., A.M. Tausch, Psychologie van opvoeding en onderwijs, Van Loghum Slaterus, Arnhem, 1967.
20. Veen, P, Meebeslissen, een veldexperiment in een hockeyclub, Van Gorcum en Comp., Assen, 1969.
21. Wallen, N.E., R.M.W. Travers, Analysis and investigation of teaching methods, in: N.L. Gage, Handbook of research on teaching, Rand McNally, Chicago, 1963, 448-505.
22. Wilde, G.J.S., Neurotische labiliteit, gemeten volgens de vragenlijstmethode, Amsterdam, 2e druk, F. van Rossen, 1970.
23. Willems, E.J., G. Beunen, M. Ostyn, R. Renson, J. Simons, P. Swalus, D. van Gerven, De differentiatie van krachtfactoren bij 12 tot 19-jarige jongens, Hermes, 6, 1971, 89-100.
24. Wijn, J.F. de, R. Luyken, Somatometrische gegevens van Nederlandse adolescenten, Centraal Instituut voor Voedingsonderzoek, T.N.O., Zeist, Rapport nr. R. 2646, maart, 1968.

SAMENVATTING EN CONCLUSIES

Ondanks het feit dat de experimenten naar het effect van extra lessen lichamelijke oefening zeer talrijk zijn geweest (schema 1.1.), kan men toch stellen dat er nauwelijks sprake is geweest van een systematische herhaling. Het leek ons daarom gewenst het eigen vooronderzoek (Oud e.a., 1971) te herhalen om eerder gevonden resultaten te kunnen controleren.

De proefpersonen in dit onderzoek waren jongens uit de vier brugklassen van het St. Ignatius College, een school voor V.W.O. en H.A.V.O. te Amsterdam. Van de 82 jongens bleven er voor het onderzoek ten slotte 70 over (tabel 2.1.). De kalenderleeftijd aan het begin van het schooljaar 1971/1972 was 12,5 jaar ($\pm$ 0,4). De vier leraren lichamelijke oefening kregen allen een brugklas toegewezen, waarvoor zij tevens als klasseleraar optraden.

Als onafhankelijke variabele werd de frequentie van het aantal lessen lichamelijke oefening per week gehanteerd. Welke twee klassen als

experimentele en welke als controlegroep zouden gaan functioneren, werd vooraf door het lot bepaald. Per week werden aan de controlegroep de normale drie lessen lichamelijke oefening gegeven. De experimentele groep kreeg in totaal vijf lessen lichamelijke oefening per week.

De twee extra lessen lichamelijke oefening werden aan het rooster toegevoegd zonder dat daarvoor andere lessen kwamen te vervallen. Hierdoor kwam het totaal aantal lessen per week voor de experimentele groep op 34 in plaats van 32.

Als uitgangspunt voor dit onderzoek dienden de doelstellingen geformuleerd in het Voorstel leerplan Rijksscholen (1968). Tevens werd dit Voorstel leerplan Rijksscholen gebruikt als richtlijn voor het jaarplan lichamelijke oefening in de brugklassen. Het jaarplan bestond uit een leerplan met eindvormen en een testplan met objectieve prestatietests (zie bijlage II). Om een zo groot mogelijke uniformiteit van les en lesindeling over de vier klassen te waarborgen, waren alle te geven lessen door de leraren tevoren opgesteld en beproefd. Vervolgens werden de lessen in een kaartsysteem ondergebracht.

De twee extra lessen lichamelijke oefening vormden uitsluitend een uitbreiding van het totale aantal lessen lichamelijke oefening in kwantitatieve zin.

Als afhankelijke variabelen werden aan het begin van het schooljaar (voortest) en aan het einde van het schooljaar (natest) de volgende groepen variabelen gemeten:

- antropometrische variabelen voor lichaamsbouw en lichaamssamenstelling (tabel 9.5.);
- fysiologische variabelen (tabel 9.7.) voor het aerobe vermogen (W_{170}, 12 min. loop), eigenschappen van het neuromusculaire systeem (kracht, snelheid en lenigheid) en de prestatie van de ventilatie (FEV%);
- gymnologische variabelen voor de vorderingen in de lichamelijke oefening (tabel 8.2.);
- psychologische variabelen voor het meten van onder andere groepscohesie (SAGS), (tabel 9.10.), neurotische labiliteit (ABV), (tabel

9.12.), prestatiemotivatie (PMT-K), attituden (SSD), (tabel 9.16.en bijlage IV), schoolattitude (tabel 9.17.), en gymnastiek voorkeur en rangnummer (tabel 9.18.).

Aan de hand van literatuuronderzoek werden ten aanzien van de te meten variabelen verwachtingen uitgesproken over het effect van de onafhankelijke variabele.

Hoewel een verhoging van drie lessen naar vijf lessen lichamelijke oefening per week op het eerste gezicht toereikend lijkt voor het aantonen van effecten van extra lessen lichamelijke oefening, doen zich enige aspecten voor die het eventuele effect kunnen maskeren, en daarom storende variabelen worden genoemd.

Als storende variabelen beschouwden wij:

1. de biologische leeftijd welke gemeten werd als de skeletleeftijd aan de hand van een röntgenfoto van hand en pols (fig. 5.2. en tabel 5.3.);
2. de buitenschoolse lichamelijke activiteit, gemeten met behulp van stappentellers (foto 5.9. en tabel 5.10.), en een vragenlijst (tabel 5.11. en bijlage V);
3. het lesgeefgedrag van de leraren lichamelijke oefening, dat werd beoordeeld met behulp van een beoordeling door deskundigen op vaktechnische aspecten (schema 5.12.) en op sociaalemotionele aspecten (schema 5.15.), en ten slotte door een beoordeling door de leerlingen zelf (schema 5.21.).

Het onderzoek had ten doel om door middel van statistische toetsing van hypothesen na te gaan of er systematische verschillen tussen de experimentele en controlegroep bestonden onder invloed van de onafhankelijke variabele.

Daarbij werd de volgende procedure gevolgd:

a. de afhankelijke variabelen werden op theoretische gronden ingedeeld in een aantal min of meer conceptueel onafhankelijke groepen;
b. daarna werd op de voortestgegevens van dit gegroepeerde materiaal door middel van factoranalyse de onderlinge relatie onderzocht bij zowel de experimentele als de controlegroep. Bij het formuleren van de hypothesen werd vervolgens gebruikgemaakt van representanten

die bij beide groepen naar voren waren gekomen;

c. bij elke op grond van de factoranalyse verkregen representant werd voor iedere hypothese gespecificeerd welke storende variabelen, en bovendien welke interacties naast de onafhankelijke variabele konden bijdragen tot de verschilscores op deze representant tussen de experimentele en controlegroep (tabel 7.7.).

Door middel van covariantieanalyse werd een vergelijking uitgevoerd tussen de gemiddelde verschilscores van de experimentele en controlegroep, waarbij rekening gehouden werd met de invloed van de storende variabelen. De storende variabele, het lesgeefgedrag van de leraren lichamelijke oefening, die niet aan de proefpersonen zelf is gemeten, kon derhalve niet in deze covariantieanalyse worden opgenomen.

Bij toetsing van de twaalf hypothesen met behulp van de covariantieanalyse (tabel 8.1.) bleek alleen de explosieve armkracht (handgrip) significant ($p \leq 0,01$) te zijn toegenomen bij de experimentele groep ten opzichte van de controlegroep, als gevolg van de extra lessen lichamelijke oefening.

Ter evaluatie van het onderwijs in de lichamelijke oefening in engere zin werd op vier tijdstippen een reeks gymnologische tests afgenomen. Analyse van deze gegevens resulteerde in een significant hogere score ($p \leq 0,009$) van de experimentele groep ten opzichte van de controlegroep (fig. 8.3.).

Voor groepjes variabelen, waarbij de factoranalyse geen gelijkluidende structuur in experimentele en controlegroep op de voortest opleverde, werd besloten tot een ongereduceerde multivariate analyse.

Bij de attitudes ten opzichte van de tien SSD concepten konden geen significante verschillen worden aangetoond als gevolg van de twee extra lessen lichamelijke oefening; voor de concepten hobby's en sport betekende dit resultaat dat het vooronderzoek niet kon worden bevestigd.

Ten aanzien van de variabelen SAGS, PMT-K en gymnastiek voorkeur en rangnummer, bleek deze groep variabelen als geheel significant te verschillen tussen experimentele en controlegroep. Bij nadere

analyse van de verschilscores lijkt vooral de ontvangen (O) en gegeven (G) score van de SAGS het verschil te veroorzaken. Wellicht kunnen deze scores die iets over het onderlinge groepsverband van de klas zeggen, als een klaseffect worden geduid.

Omdat het klas- en/of leraareffect in het onderzoek niet direct in de covariantieanalyse kon worden opgenomen, is de eventuele invloed van deze storende variabele op exploratieve wijze alsnog onderzocht (schema 9.1.). Voor elke van de twaalf representanten werd de ratio berekend van het effect van de onafhankelijke variabele plus het klas- en/of leraareffect tussen experimentele en controlegroep enerzijds en het klas- en/of leraareffect binnen de experimentele en controlegroep anderzijds.

Bij een aantal variabelen was deze ratio klein en er mag worden verondersteld dat het verschil tussen experimentele en controlegroep voor een niet onbelangrijk deel op rekening van een klas- en/of leraareffect zou kunnen komen. Bij de explosieve armkracht was deze ratio zeer groot, zodat er geen enkele aanleiding bestaat om een klas- en/of leraareffect te veronderstellen, zodat het in de covariantieanalyse gevonden resultaat bij deze variabele nog versterkt wordt.

Het bestaan van een leraareffect wordt verder nog gesteund door gegevens uit de verschillende beoordelingssystemen van het lesgeefgedrag, waaruit blijkt dat vooral de leraren van de experimentele groep in een aantal opzichten van elkaar verschillen. Op grond van de resultaten van de covariantieanalyse is op statistische gronden aanleiding om te veronderstellen dat de invloeden van twee extra lessen lichamelijke oefening, de biologische leeftijd, en de buitenschoolse lichamelijke activiteit op een groot aantal afhankelijke variabelen van ondergeschikte betekenis zijn geweest. Het blijft vooralsnog de vraag of:

- de twee extra lessen lichamelijke oefening per week een te geringe verhoging zijn geweest;
- de juiste afhankelijke en storende variabelen zijn gekozen;
- de betrouwbaarheid van de storende variabelen te wensen overliet;
- een mogelijke invloed van andere, ons onbekende storende variabelen een rol heeft gespeeld.

Samenvattend kan gesteld worden dat het effect in zijn algemeenheid van twee extra lessen lichamelijke oefening per week gedurende een schooljaar op de lichamelijke en geestelijke ontwikkeling van 12 en 13-jarige jongens in de brugklas niet kon worden aangetoond. Echter de vorderingen in de les lichamelijke oefening blijken in dit onderzoek, evenals in het vooronderzoek, zeer duidelijk toe te nemen bij de experimentele groep ten opzichte van de controlegroep.

Hiermee wordt echter geen enkele uitspraak gedaan over het belang van het normale aantal lessen lichamelijke oefening per week voor de brugklassen. Dit moge blijken uit de waardering van het vak lichamelijke oefening ten opzichte van de andere vakken. Zowel aan het begin als aan het einde van het schooljaar wordt het vak lichamelijke oefening door de leerlingen gemiddeld op de tweede tot derde plaats gekozen van de in totaal dertien vakken. Bovendien blijkt deze hoge waardering tijdens het schooljaar bij zowel de experimentele als controlegroep zelfs nog iets toe te nemen, terwijl de gemiddelde schoolattitude daalt.

BIJLAGEN

I Copie van een brief van het Ministerie van Onderwijs en Wetenschappen

II Jaarplan lichamelijke oefening op de brugklassen van het St. Ignatius College

III Volledige beschrijving van de eenvoudige prestatietests

IV Standaard semantische differentiaal

V Vragenformulier buitenschoolse activiteiten en rookgewoonten

MINISTERIE VAN ONDERWIJS EN WETENSCHAPPEN
NIEUWE UITLEG 1 - 'S-GRAVENHAGE

Kenmerk	Telefoon (070) - 74	Datum
AVO/VHO - 501361 I	1960	24 april 1972.

AAN het bestuur van de St. Willebrordus-stichting,

Hobbemakade 51,

Amsterdam.

Uw brief van 9 augustus 1971.

Onderwerp afwijking lessentabel brugjaar i.v.m. onderzoek invloed extra lessen lichamelijke oefening.

Met het gestelde in bovenvermelde brief ga ik akkoord.

De minister/staatssecretaris,
namens deze,
de chef van de afdeling,
voor deze,

(drs. H. de Vries)

PERIODE 1	LEERPLAN
Eindvormen:	Materiaal:
1. Duikelen voor- en achterover	Ringen
2. Bovenhandse strekworp	Handbal
3. 3-pas ritme met tippen	Handbal
4. Rollen voorover	Mat
5. Rollen achterover	Mat

PERIODE 2	LEERPLAN
6. Dribbelen	Basketbal
7. Pivoteren	Basketbal
8. Handstand	'Verticaal vlak'
9. Duikelen voor/achterover	Brug ongelijk
10. Spreidsprong	Bok
11. Klimmen	Touw

PERIODE 1 — TESTPLAN

Testnummer : Testbeschrijving

1. Ringen stil op schouderhoogte: vanuit buiklig op de mat komen tot zit in de ringen, daarna 1 x voorover duikelen en weer komen tot buiklig op de mat:
   1 p. gehaald binnen 30 sec.
   2 p. gehaald binnen 15 sec.
   3 p. gehaald binnen 11 sec.
   4 p. gehaald binnen 8 sec.

2. (extra test) Schuine bank/kastlengte/schuine bank: in zo weinig mogelijk passen erover. Elke stap wordt geteld:
   1 p. 6 stappen
   2 p. 5 stappen
   3 p. 4 stappen
   4 p. 3 stappen

3. Matten: rollen voorover gevolgd door rollen achterover van stand tot stand:
   1 p. gehaald met extra steun van de handen (voorover)
   2 p. gehaald met extra steun van de knieën en/of elleboog (achterover)
   3 p. gehaald met voeten verplaatsen bij het komen tot stand (voorw./achterw.)
   4 p. gehaald zonder genoemde fouten

PERIODE 2 — TESTPLAN

4. Basketbal dribbelen: de bal na elke stuit aanraken en zowel bal als één voet moeten de lijnen gepasseerd zijn alvorens te keren:
   1 p. gehaald binnen 21 sec.
   2 p. gehaald binnen 19 sec.
   3 p. gehaald binnen 17 sec.
   4 p. gehaald binnen 16 sec.

5. Handstand tegen verticaal vlak in 4 - 8 pogingen achter elkaar, met minstens 1 voet verticaal vlak raken:
   1 p. 1 x gehaald (of 2 x met 1 voet)
   2 p. 2 x gehaald (of 4 x met 1 voet)
   3 p. 3 x gehaald (of 6 x met 1 voet)
   4 p. 4 x gehaald (of 8 x met 1 voet)

6. Bok (breedte) op schouderhoogte spreidsprong:
   1 p. met afzet vanaf bank
   2 p. met afzet vanaf de grond
   3 p. met afzet vanaf de grond vóór de bank
   4 p. als 3 en met gesloten voeten landen

7. Touwklimmen vanuit stand, plafond aantikken en tegelijk 'ja' roepen:
   1 p. gehaald binnen 30 sec.
   2 p. gehaald binnen 20 sec.
   3 p. gehaald binnen 15 sec.
   4 p. gehaald binnen 10 sec.

PERIODE 3	LEERPLAN
Eindvormen:	Materiaal:
12. Zwaaien in strekhang	Ringen
13. Duikelen voor- en achterover	Rekstok
14. Dribbelen en met 2-pas ritme tot stilstand komen	Basketbal
15. Vouwhang	Ringen
16. Zweefrol	Mat
17. Buighang	Ringen

PERIODE 4	LEERPLAN
18. Vangen en werpen	Basketbal
19. Borstwaartsom	Brug ongelijk
20. Lay-up shot met ophouden van bal	Basketbal
21. Ondersprong	Brug ongelijk
22. Handstand doorrollen	Mat
23. Zwaaien met afsprong achter	Ringen

PERIODE 3 — TESTPLAN

Testnummer:	Testbeschrijving
8.	Rekstok op schouderhoogte: voorover duikelen gevolgd door achterover duikelen onder rekstok en tussen handen door: 1 p. gehaald 2 p. gehaald zonder de grond tussentijds met voeten aan te raken 3 p. gehaald zonder rekstok met voeten aan te raken 4 p. gehaald met gesloten voeten en knieen
9.	Zweefrol voorover op mat en komen tot stand zonder hulp van de handen: 1 p. met aanloop en sprongetje op bank 2 p. met aanloop en afzet vóór bank 3 p. met aanloop en afzet voor 2 kastdelen 4 p. als 3 p. gevolgd door ½ draai om lenteas en rol achterover
10.	Ringen op reikhoogte: vanuit vouwhang tot buighang komen (vouwhang gestrekt aangeven): 1 p. 3 x achtereen gehaald 2 p. 8 x achtereen gehaald 3 p. 14 x achtereen gehaald 4 p. 20 x achtereen gehaald
11. (extra test)	Shuttle run met behulp van volleyballijnen (elke lijn met minstens één voet passeren): 1 p. binnen 18 sec. 2 p. binnen 17 sec. 3 p. binnen 16 sec. 4 p. binnen 15 sec.

PERIODE 4 — TESTPLAN

12.	Brug ongelijk (1 legger op borsthoogte, andere legger op halve reikhoogte): borstwaarts omtrekken: 1 p. gehaald om de lage legger (mbv. de hoge) 2 p. gehaald om de hoge legger (mbv. de lage) 3 p. gehaald om de lage legger (gezicht buitenw.) 4 p. gehaald om de hoge legger zonder hulp van de lage (gezicht binnen- of buitenwaarts)
13.	Lay-up shot met de bal opgehouden door 11.: gedurende 1 min. lay-ups maken. Degene die schiet moet via de middencirkel teruglopen. Aantal doelpunten wordt geteld: 1 p. meer dan 2 doelpunten 2 p. meer dan 4 doelpunten 3 p. meer dan 6 doelpunten 4 p. meer dan 8 doelpunten
14.	In 20 sec. zo vaak mogelijk komen tot vluchtige handstand met minstens 1 voet telkens verticaal vlak raken: 1 p. meer dan 1 x 2 p. meer dan 5 x 3 p. meer dan 10 x 4 p. meer dan 15 x

LEERPLAN

Eindvormen:	Materiaal:

PERIODE 5	LEERPLAN	
24.	Radslag met steun op verhoogd vlak	Mat en bank
25.	Hurksprong	Paard met beugels
26.	Hurksprong	Kast
27.	Verplaatsen in hang: voorwaarts, achterwaarts en zijwaarts	Ladder horizontaal
28.	Zweefrol	Kastlengte
29.	Handstand	'Verhoogd vlak' (bank)
30.	Zwaaien en klimmen	Touw

PERIODE 6	LEERPLAN	
31.	Lay-up shot buiten de zône	Basketbal
32.	Ondersprong	Rekstok
33.	Buikdraai	Rekstok
34.	Zwaaien met halve draai om le-as	Bok

TESTPLAN

Testnummer:	Testbeschrijving:
15. (extra test)	Met basketbal dribbelen gedurende 30 sec. tussen twee lijnen op zes m. afstand van elkaar (telkens zowel 1 voet als bal over de lijn): 1 p. 10 x gehaald 2 p. 11 x gehaald 3 p. 12 x gehaald 4 p. 13 x gehaald

PERIODE 5

TESTPLAN

16. Hurksprong vanaf lattenplankje (één na hoogste stand) over kastbreedte of vanaf reutherplankje over paardbreedte:
1 p. gehaald over paard met steun van beugels
2 p. gehaald over paard zonder steun van beugels
3 p. gehaald over kast
4 p. gehaald over kast met med. bal-afstand tussen toestel en afzetplankje

17. Rollen voorover op de lengte van de kast:
1 p. gehaald op kast minus 1 deel
2 p. idem, met gestrekte knieën en gesloten voeten
3 p. gehaald op volledige kast
4 p. idem, met gestrekte knieën en gesloten voeten

18. Handstand op verhoogd vlak met tegenhouden van medell.:
1 p. gehaald zonder verhoogd vlak met hulp
2 p. gehaald zonder verhoogd vlak zonder hulp
3 p. gehaald op verhoogd vlak met hulp
4 p. gehaald op verhoogd vlak zonder hulp

19. Touwklimmen (op en neer): 'kaats' op de grond en met hoofd plafond raken:
1 p. 10 m. totaal geklommen
2 p.. 15 m. totaal geklommen
3 p. 20 m. totaal geklommen
4 p. 25 m. totaal geklommen

PERIODE 6

TESTPLAN

LEERPLAN

Eindvormen:		Materiaal
35.	Spreidsprong met halve draai om le-as	Bok
36.	Sprong worp	Handbal
37.	Kipslag tot brugstand	Mat

ZWEMMEN (ALLE PERIODES) LEERPLAN

Afhankelijk van de reeds aanwezige vaardigheid, beoefening van verschillende zwemslagen, het starten, springen en duiken, het keren en het watertrappen, alsmede het vervoeren van een medeleerling, desgewenst gericht op het afleggen van de examens ter verkrijging van schoolzwemdiploma's.

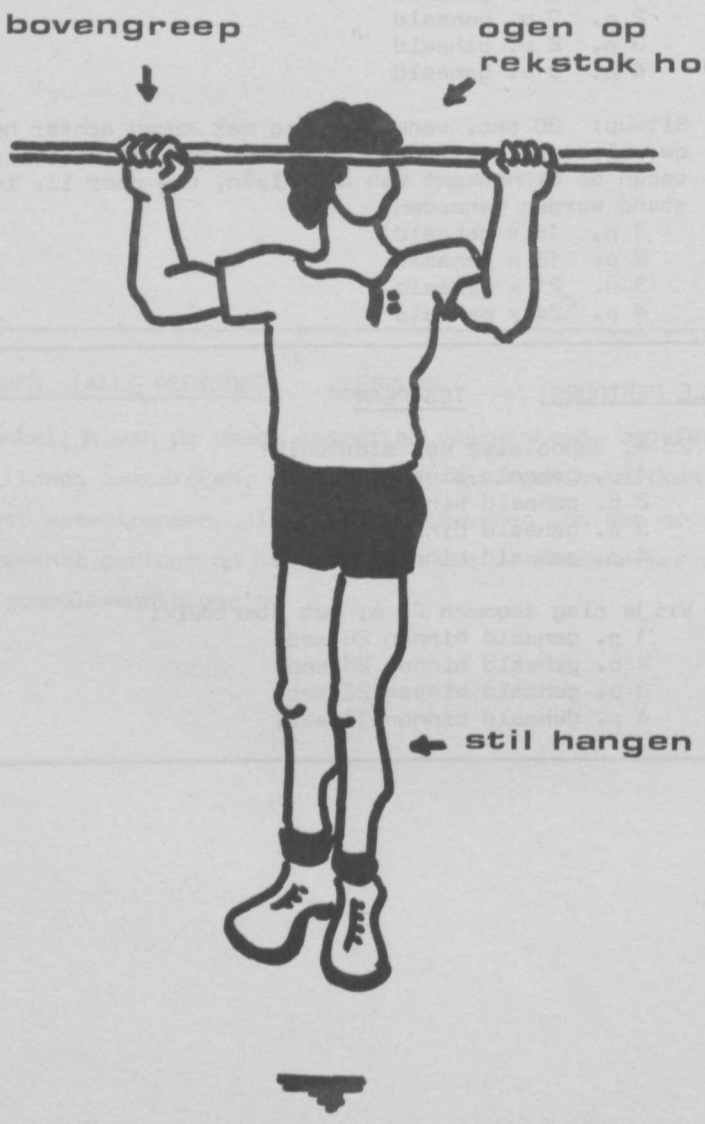
TESTPLAN

Testnummer:	Testbeschrijving:
21.	Spreidsprong (boklengte, afzet zonder plankje): 1 p. gehaald over bok in laagste stand 2 p. gehaald over bok op tepelhoogte 3 p. idem, met steun van twee handen over de helft 4 p. idem, met neerkomen op tweede mat (breedte)
22. (extra test)	Hinkstapsprong met afzet vanaf reutherplank en landing op de mat: 1 p. 6 m. gehaald 2 p. 7 m. gehaald 3 p. 8 m. gehaald 4 p. 9 m. gehaald
23. (extra test)	Sit-up: 30 sec. vanuit ruglig met armen achter hoofd en de ellebogen tegen de grond komen tot zit met ellebogen tegen de binnenkant van de knieën, die door ll. in buigstand worden gehouden: 1 p. 15 x gehaald 2 p. 18 x gehaald 3 p. 21 x gehaald 4 p. 24 x gehaald

ZWEMMEN (ALLE PERIODES) TESTPLAN

1.	25 m. schoolslag met startduik: 1 p. gehaald binnen 30 sec. 2 p. gehaald binnen 27 sec. 3 p. gehaald binnen 25 sec. 4 p. gehaald binnen 23 sec.
2.	Vrije slag zwemmen 25 m. met startduik: 1 p. gehaald binnen 28 sec. 2 p. gehaald binnen 25 sec. 3 p. gehaald binnen 22 sec. 4 p. gehaald binnen 19 sec.

BENT ARM HANG



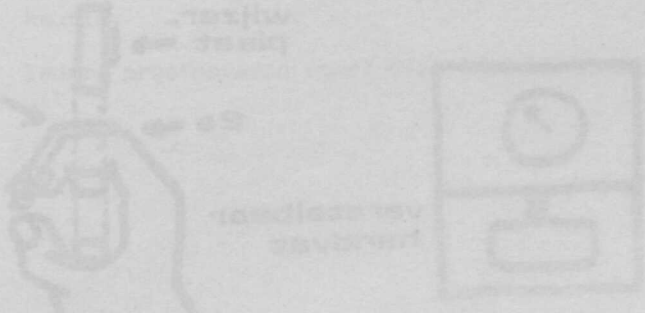
TEST : BENT ARM HANG

OPSTELLING: Rekstok op reikhoogte. Indien mogelijk, meerdere rekstokken op verschillende hoogten klaar. Een tester met twee stopwatches.

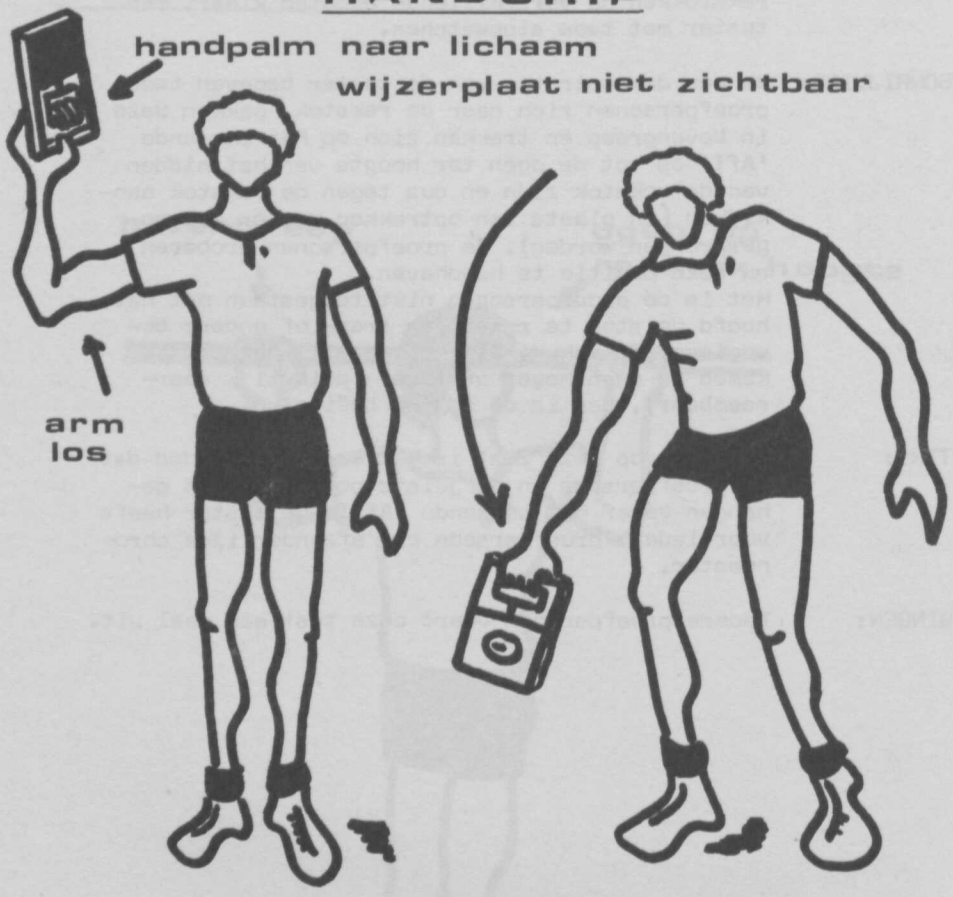
BESCHRIJVING: Na een demonstratie van de tester begeven twee proefpersonen zich naar de rekstok, pakken deze in bovengreep en trekken zich op het commando 'Af!' op tot de ogen ter hoogte van het midden van de rekstok zijn en dus tegen de rekstok aankijken (in plaats van optrekken mag er ook opgesprongen worden). De proefpersonen proberen nu deze positie te handhaven.
Het is de proefpersonen niet toegestaan met het hoofd de stok te raken, en trap- of andere bewegingen te maken. Zij moeten stil hangen.
Komen de ogen hoger of lager (duidelijk waarneembaar), dan is de poging beëindigd.

METING: De score op deze test is het aantal seconden dat de proefpersoon in de juiste positie heeft gehangen vanaf het commando 'Af!'. De tester heeft voor iedere proefpersoon een afzonderlijke chronometer.

POGINGEN: Iedere proefpersoon voert deze test een maal uit.

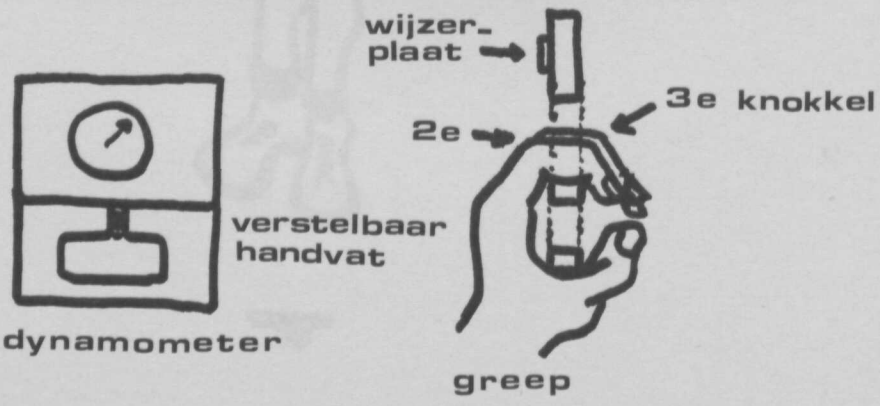


HANDGRIP



UITGANGSHOUDING

KNIJP AKTIE



TEST : HANDGRIP

OPSTELLING: Er wordt gebruikgemaakt van de handdynamometer (type Bettendorf). Een tester.

BESCHRIJVING: De dynamometer wordt in de palm van de voorkeurhand genomen.
De wijzerplaat moet van de handpalm afgekeerd zijn. De vingers zijn gebogen over de bovenkant van het handvat. Het handvat wordt zo gesteld dat het gedeelte van de vingers tussen tweede en derde knokkel het handvat omklemmen.
De proefpersoon staat.
De arm met de dynamometer is gebogen en geheven. De hand bevindt zich ter hoogte van de schouders. De arm moet los van het lichaam zijn. De handpalm is naar het lichaam gekeerd.
Op het commando 'Knijp!' zwaait de arm met dynamometer naar beneden terwijl de proefpersoon fel en zo hard mogelijk knijpt.
De tester moet een demonstratie geven van greep en positie van de arm.
Hij moet erop toezien dat tijdens de testpogingen de proefpersoon met geen enkel deel van zijn arm, noch met de dynamometer zijn lichaam raakt.
De tester moet gedurende de eerste poging elke onjuiste gang van zaken verbeteren.
Wanneer de regels zijn overschreden moet het resultaat niet geteld worden, maar wel de poging.
De noodzaak van een korte, felle kneep wordt benadrukt. Indien de proefpersoon langzaam begint te knijpen zodra hij de dynamometer heeft, zal het zijn score verminderen als gevolg van spiervermoeidheid.
Iedere proefpersoon krijgt twee pogingen gescheiden door tenminste een volle minuut rust. Zonder de rust is het waarschijnlijk dat het resultaat bij elke nieuwe poging zal verminderen.

METING: Het resultaat wordt afgelezen tot op 1 kg. nauwkeurig.

POGINGEN: Iedere proefpersoon voert deze test twee maal uit.

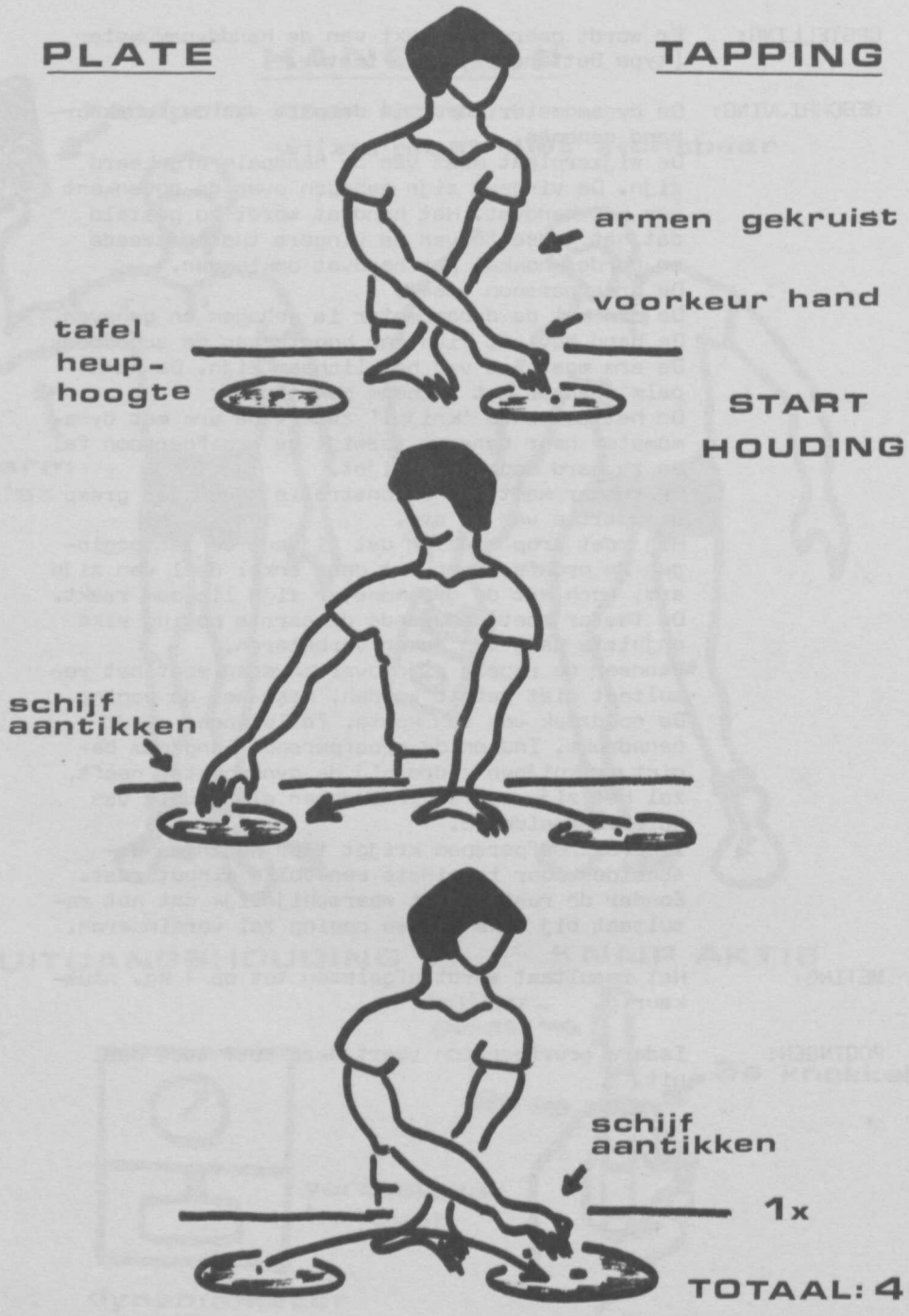


PLATE
TAPPING
armen gekruist
voorkeur hand
tafel
heup-
hoogte
START
HOUDING
schijf
aantikken
schijf
aantikken
1x
TOTAAL: 40x

TEST: PLATE TAPPING

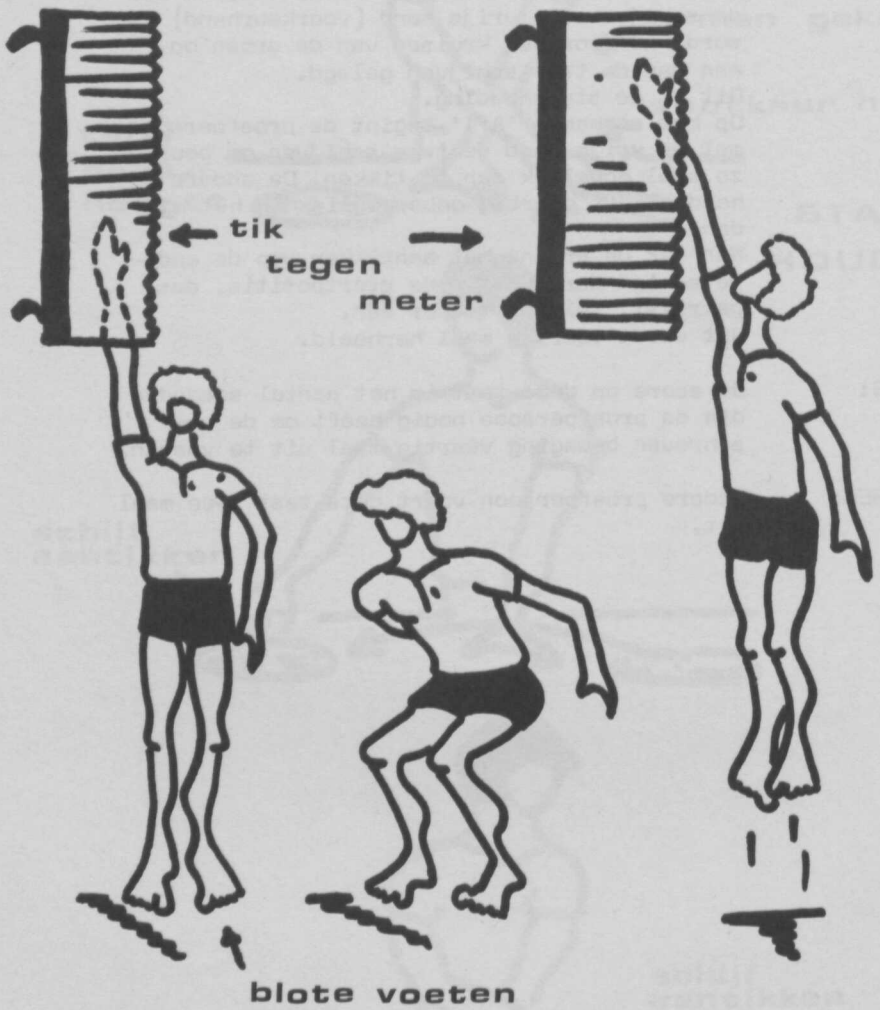
OPSTELLING: Plate tapping tafel op heuphoogte ingesteld. Tester met een stopwatch.

BESCHRIJVING: Na een demonstratie door de tester wordt de tafel op heuphoogte van de proefpersoon ingesteld. Deze houdt zijn voorkeurhand vrij en legt de andere hand in het midden tussen beide schijven. De vrije hand (voorkeurhand) wordt nu door het kruisen van de armen op een van de twee schijven gelegd.
Dit is de starthouding.
Op het commando 'Af!' begint de proefpersoon met de vrije hand de twee schijven om beurten zo snel mogelijk aan te tikken. De andere hand blijft daarbij onbeweeglijk in het midden.
Wanneer de arm na het aantikken van de andere schijf terug is in de startpositie, dus gekruist, telt de tester één.
Dit wordt veertig maal herhaald.

METING: De score op deze test is het aantal seconden die de proefpersoon nodig heeft om de beschreven beweging veertig maal uit te voeren.

POGINGEN: Iedere proefpersoon voert deze test twee maal uit.

VERTICAL JUMP



REIKHOOGTE SPRONGHOOGTE

TEST: VERTICAL JUMP

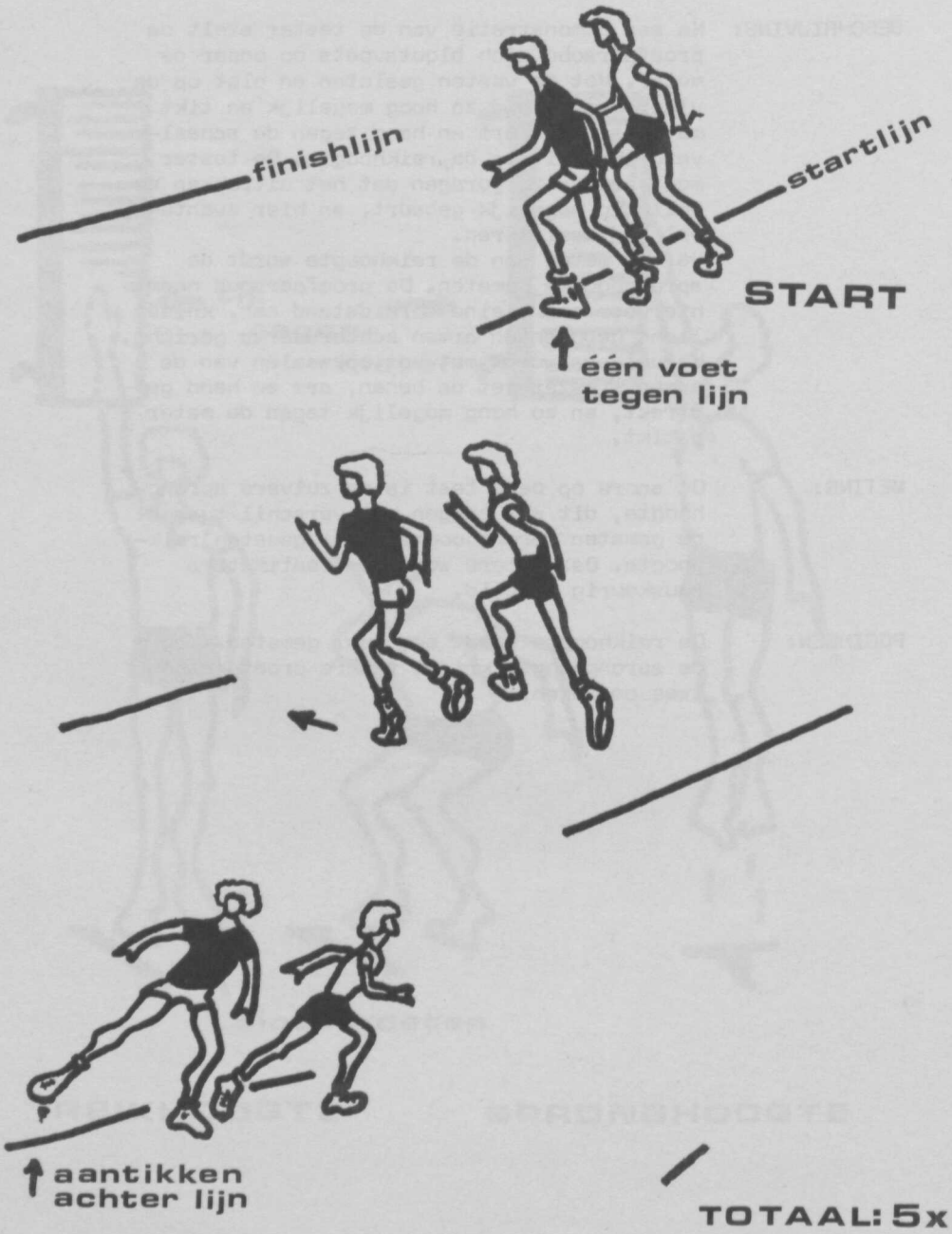
OPSTELLING: Spronghoogtemeter aan een muur of wandrek bevestigd. Proefpersonen op blote voeten. Een tester.

BESCHRIJVING: Na een demonstratie van de tester stelt de proefpersoon zich blootsvoets op onder de meter. Met de voeten gesloten en plat op de vloer, reikt hij zo hoog mogelijk en tikt met gestrekte arm en hand tegen de schaalverdeler. Dit is de reikhoogte. De tester moet ervoor zorgdragen dat het uitrekken zo volledig mogelijk gebeurt, en hier eventueel bij assisteren.
Na het meten van de reikhoogte wordt de spronghoogte gemeten. De proefpersoon neemt hiervoor een kleine spreidstand aan, knieën licht gebogen en armen achterwaarts gericht. Vervolgens wordt met het opzwaaien van de armen afgezet met de benen, arm en hand gestrekt, en zo hoog mogelijk tegen de meter getikt.

METING: De score op deze test is de zuivere spronghoogte, dit wil zeggen het verschil tussen de gemeten spronghoogte en de gemeten reikhoogte. Deze score wordt in centimeters nauwkeurig bepaald.

POGINGEN: De reikhoogte wordt een maal gemeten. Voor de spronghoogte krijgt iedere proefpersoon twee pogingen.

50 meter SHUTTLE RUN



TEST: 50 m. SHUTTLE RUN

OPSTELLING: Twee evenwijdige lijnen, tien meter uit elkaar. Twee proefpersonen met gymnastiekschoenen aan. Een tester aan de finishlijn met twee stopwatches.

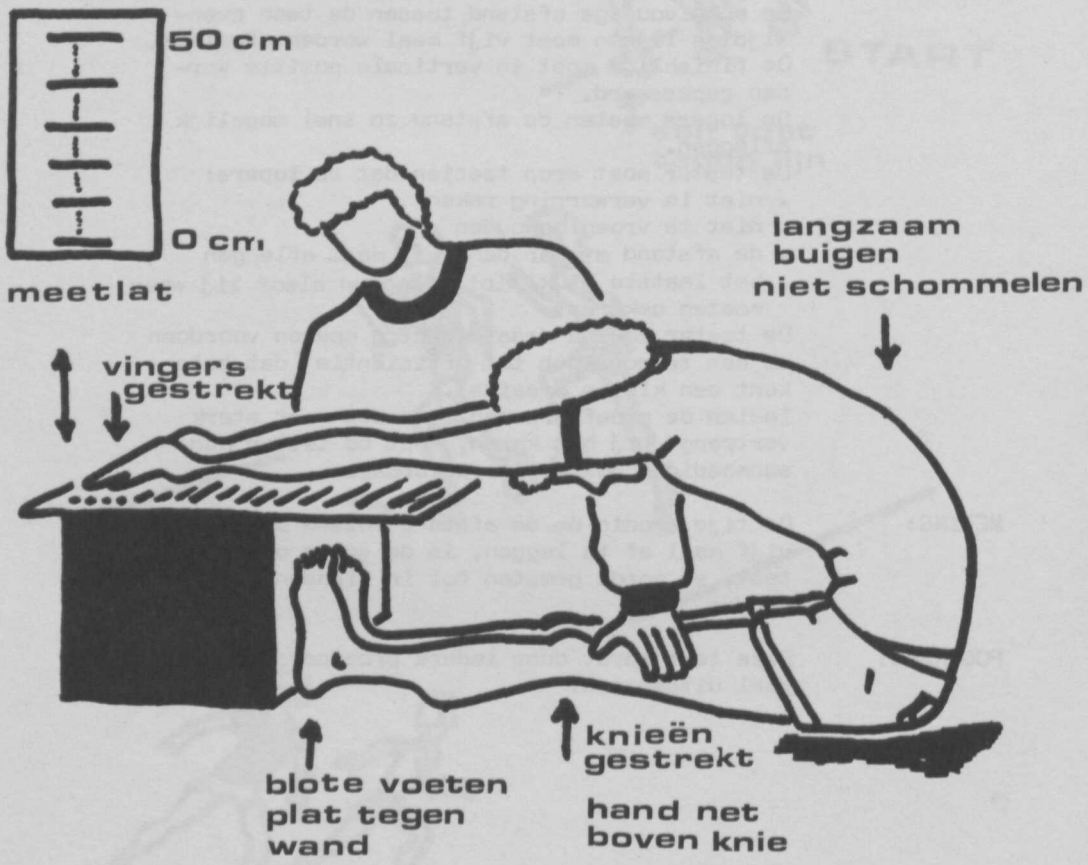
BESCHRIJVING: De proefpersonen staan achter de startlijn, met één teen tegen de lijn; ze nemen de staande starthouding aan.
Op het commando 'Af!' rennen de proefpersonen naar de tegenoverliggende lijn, raken de vloer achter de lijn met een der voeten aan, keren terug naar de startlijn, raken de vloer achter de startlijn met één voet aan, en vice versa.
De enkelvoudige afstand tussen de twee evenwijdige lijnen moet vijf maal worden afgelegd.
De finishlijn moet in verticale positie worden gepasseerd.
De lopers moeten de afstand zo snel mogelijk afleggen.
De tester moet erop toezien dat de lopers:
- niet in verwarring raken
- niet te vroeg ophouden
- de afstand minder dan vijf maal afleggen
- het laatste stuk niet afleggen alsof zij weer moeten omkeren.

De tester zou de draaibeweging moeten voordoen om aan te moedigen tot efficiëntie (dat betekent een kleine draaias).
Indien de proefpersonen iets doen wat sterk vertraagt bij het keren, moet de tester hen aanmoedigen om sneller te draaien.

METING: De tijd, nodig om de afstand tussen de lijnen vijf maal af te leggen, is de score op deze test, en wordt gemeten tot in tienden van seconden.

POGINGEN: Deze test wordt door iedere proefpersoon twee maal uitgevoerd.

SIT AND REACH



TEST: SIT AND REACH

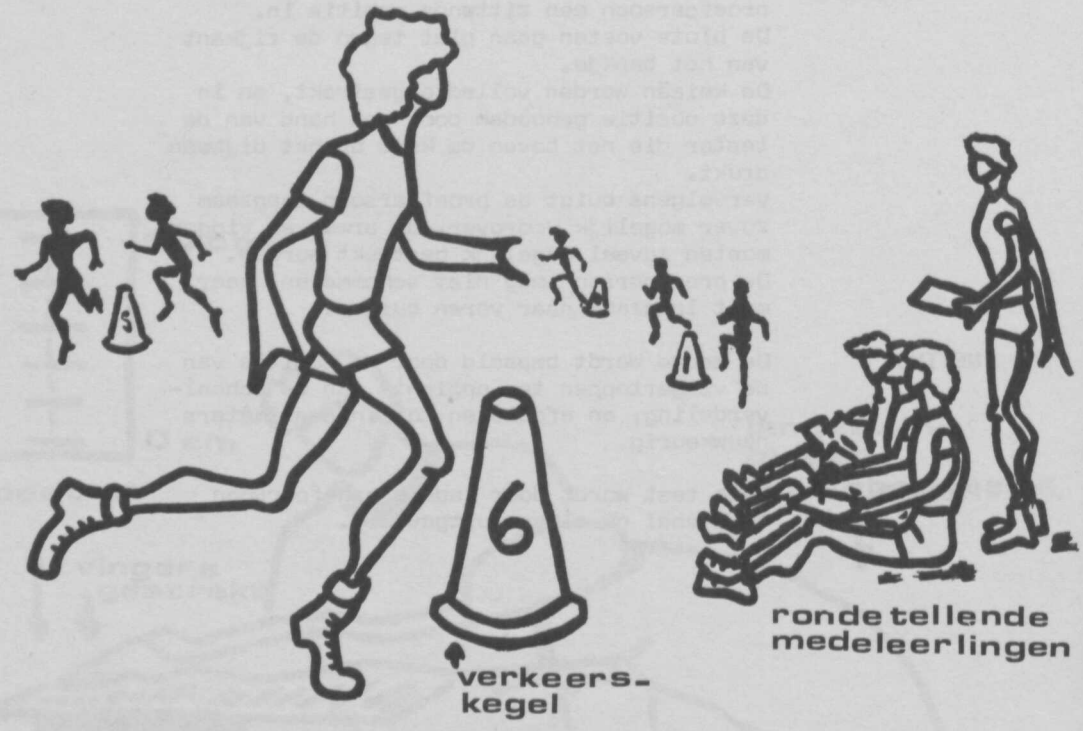
OPSTELLING: Voetenbankje met horizontaal een schaalverdeling in centimeters. Proefpersonen op blote voeten.
Een tester.

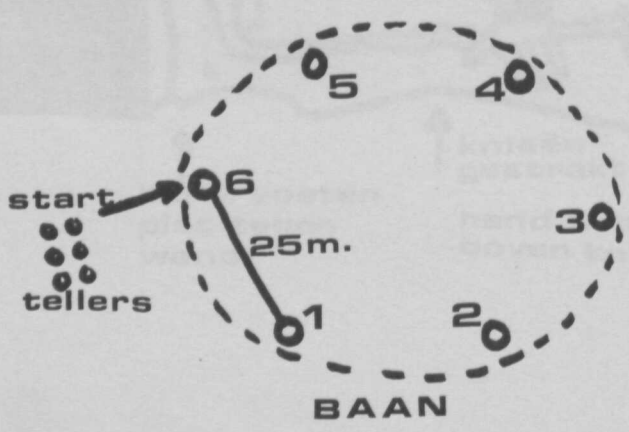
BESCHRIJVING: Na een demonstratie van de tester neemt de proefpersoon een zittende positie in.
De blote voeten gaan plat tegen de zijkant van het bankje.
De knieën worden volledig gestrekt, en in deze positie gehouden door een hand van de tester die net boven de knie op het dijbeen drukt.
Vervolgens buigt de proefpersoon langzaam zover mogelijk voorover. De armen en vingers moeten zoveel mogelijk gestrekt worden.
De proefpersoon mag niet schommelen, maar moet langzaam naar voren buigen.

METING: De score wordt bepaald door de positie van de vingertoppen ten opzichte van de schaalverdeling, en afgelezen tot in centimeters nauwkeurig.

POGINGEN: Deze test wordt door iedere proefpersoon drie maal na elkaar uitgevoerd.

12 minuten LOOP





TEST: 12 MINUTEN RUN WALK

OPSTELLING: Een terrein waarop een baan van 150 m. kan worden uitgezet. Deze baan wordt door zes kegels, genummerd van 1 tot en met 6, gemarkeerd in zes gelijke afstanden van elk 25 m.. Een tester.

BESCHRIJVING: De proefpersonen krijgen instructie om te trachten de volle twaalf minuten te blijven lopen en dus niet in een hoog tempo te starten. Zij mogen het lopen afwisselen met gaan, maar niet langer dan tot de volgende kegel. Tijdens het lopen wordt aan de startlijn de tijd afgeroepen. Het einde wordt door een fluitsignaal aangegeven. De proefpersonen onthouden zelf welk kegelnummer zich op dat moment voor hen bevindt.
De start vindt plaats bij kegel 6.
Er moet buiten de kegels om gelopen worden, zonder deze aan te raken. Het lopen mag worden afgewisseld met gaan, maar ten hoogste 25 meter achter elkaar.
De proefpersonen worden verdeeld in twe groepen. Iedere proefpersoon van de wachtende groep houdt op een formulier het aantal afgelegde ronden van een bepaalde proefpersoon van de lopende groep bij, en noteert ook het nummer van de kegel die de lopende proefpersoon bij het eindsignaal voor zich heeft.

METING: De afgelegde afstand wordt als volgt berekend en genoteerd:
voorbeeld: aantal gehele ronden is 16 en in de laatste ronde is de proefpersoon gekomen tot kegel 5 - dan is de afgelegde afstand:
16 x 150 + 5 x 25 = 2.400 + 125 = 2.525 meter.

POGINGEN: Iedere proefpersoon voert deze test een maal uit.

De Standaard semantische differentiaal (S.S.D.)

1. Werkwijze

Boven aan iedere bladzijde stond een concept vermeld met daaronder de 12 adjectief paren; deze tegengestelde adjectief paren stonden aan weerszijden van een 7-puntsschaal. Aan de proefpersonen werd gevraagd om achtereenvolgens op iedere schaal hun oordeel of mening over dat concept te geven door in een van de zeven hokjes een kruisje te zetten. Op deze wijze werden de concepten een voor een over de twaalf schalen beoordeeld.

2. Volgorde der schalen

De twaalf schalen, dit wil zeggen voor iedere factor vier schalen, zijn als volgt gerangschikt:

zwaar (heavy)	(7)	__	__	__	__	__	(1)	licht (light)
passief (passive)	(1)	__	__	__	__	__	(7)	actief (active)
prettig (pleasant)	(7)	__	__	__	__	__	(1)	naar (unpleasant)
zwak (weak)	(1)	__	__	__	__	__	(7)	sterk (strong)
snel opgewonden (exiting)	(7)	__	__	__	__	__	(1)	kalm (calming)
vlug (fast)	(7)	__	__	__	__	__	(1)	langzaam (slow)
ongelukkig (sad)	(1)	__	__	__	__	__	(7)	gelukkig (happy)
onaardig (violent)	(1)	__	__	__	__	__	(7)	aardig (gentle)
saai (dull)	(1)	__	__	__	__	__	(7)	boeiend (fascinating)
groot (large)	(7)	__	__	__	__	__	(1)	klein (small)
zacht (soft)	(1)	__	__	__	__	__	(7)	hard (hard)
slecht (bad)	(1)	__	__	__	__	__	(7)	goed (good)

3. Volgorde der concepten

De concepten werden als volgt aangeboden:

1. leren
2. doorzettingsvermogen
3. samenwerken
4. vroeg opstaan
5. hobby's
6. transpireren
7. sport
8. elkaar helpen
9. rekenen
10. inspanning

4. Scoring

De hokjes in de 7-puntsschaal werden gewaardeerd van een tot en met zeven, zodat iedere factorscore een minimum van 4 en een maximum van 28 had. Het nulpunt of neutraal punt wordt vertegenwoordigd door de waardering 4. Door middel van een sjabloon konden de formulieren worden gescoord.

Omdat de schalen wisselend van links naar rechts en omgekeerd oplopen van 1 tot 7, is van iedere schaal de waardering aangegeven.

5. Instructie bij afname

NAAM: DATUM VAN INVULLING

GEBOORTEDATUM:

KLAS:

Met deze vragenlijst willen wij graag weten wat je van verschillende dingen denkt.
Op iedere bladzijde vind je bovenaan een bepaald woord of begrip staan, waarover wij je oordeel vragen. Daaronder staan telkens een reeks tegengestelde woorden.
Hoe ga je nu je oordeel geven? Bij voorbeeld:

mijn vader

prettig [| | | | | |] onprettig

slordig [| | | | | |] netjes

etc.

Tussen elk woordenpaar staan zeven hokjes, zeven plaatsen waar je in één ervan een kruisje moet zetten.
Als je jouw vader uitgesproken prettig vindt, dan zet je het kruisje in het hokje uiterst links en dus vlak naast prettig:

prettig [x| | | | | |] onprettig

Heb je een uitgesproken voorkeur voor onprettig, als het gaat over jouw vader, dan zet je het kruisje uiterst rechts en dus vlak naast onprettig.

prettig [| | | | | |x] onprettig

Als je vindt dat jouw vader bij beide woorden evenveel thuishoort of neutraal is, dan zet je het kruisje precies in het midden, dus even ver van prettig als van onprettig.

prettig [| | |x| | |] onprettig

Je begrijpt dat er zo nog andere mogelijkheden zijn voor de plaatsing van het kruisje:
2 plaatsen naar rechts en 2 plaatsen naar links, naarmate je voorkeur meer of minder uitgaat naar rechter of linker woord boven aan de bladzijde.

Belangrijk om te onthouden voordat je gaat beginnen:

- plaats het kruisje in het hokje.
- sla geen regel over.
- plaats maar telkens één kruisje tussen één woordenpaar.
- gebruik alle hokjes en kies niet alleen het middelste of de twee uitersten.
- soms zal het moeilijk zijn om het kruisje te plaatsen; ook dan moet je niet te lang nadenken en afgaan op je eerste indruk. Werk zo snel als mogelijk door. Er zijn geen goede of foute antwoorden mogelijk en ook geen strikvragen. Je hoeft dus niet terug te kijken naar wat je eerder hebt ingevuld.

Vragenformulier met betrekking tot BUITENSCHOOLSE ACTIVITEITEN

Naam:
Klas: Datum:

Vervoer van en naar school

. Met wat voor een vervoermiddel ga je naar school?	O lopend	tijd:
	O per fiets	tijd:
	O per brommer	
	O per auto	
	O openb. vervoer (+ lopen)	tijd:
		10 × _____

1. Weektotaal :

Activiteiten in verenigingsverband

. Ben je lid van een vereniging? (sport, muziek, dansen, etc.)

. Hoeveel keer per week train je? . . . Hoe lang per keer? . . .

. Hoe vaak per maand volg je de training? . . .

. Is de training binnen of buiten? . . .

. Kan buitentraining bij slecht weer naar binnen? . . .
(trainingstijd halveren bij buiten, tenzij indoor)

trainingstijd :

. Vervoermiddel en tijd in verband met training (buiten = $\frac{1}{2}$ tijd) vervoertijd :

. Speel je wedstrijden? . . . Hoe vaak per week? . . .

. Hoe lang per wedstrijd? . . . wedstrijdtijd :

. Vervoermiddel en tijd in verband met wedstrijden (thuis) (helft van aantal wedstrijden voor vervoertijd) (buiten $-\frac{1}{4}$) vervoertijd : _____

2. Weektotaal :

Activiteiten buiten de verenigingen

. Wat doe je na schooltijd? (boodschap, hond, krant, buitenspelen, etc.)

. op maandag (keren per maand)	tijd :
. op dinsdag	tijd :
. op woensdag (vrije middag)	tijd :
. op donderdag	tijd :
. op vrijdag	tijd :
. op zaterdag	tijd :
. op zondag	tijd : _____

3. Weektotaal :