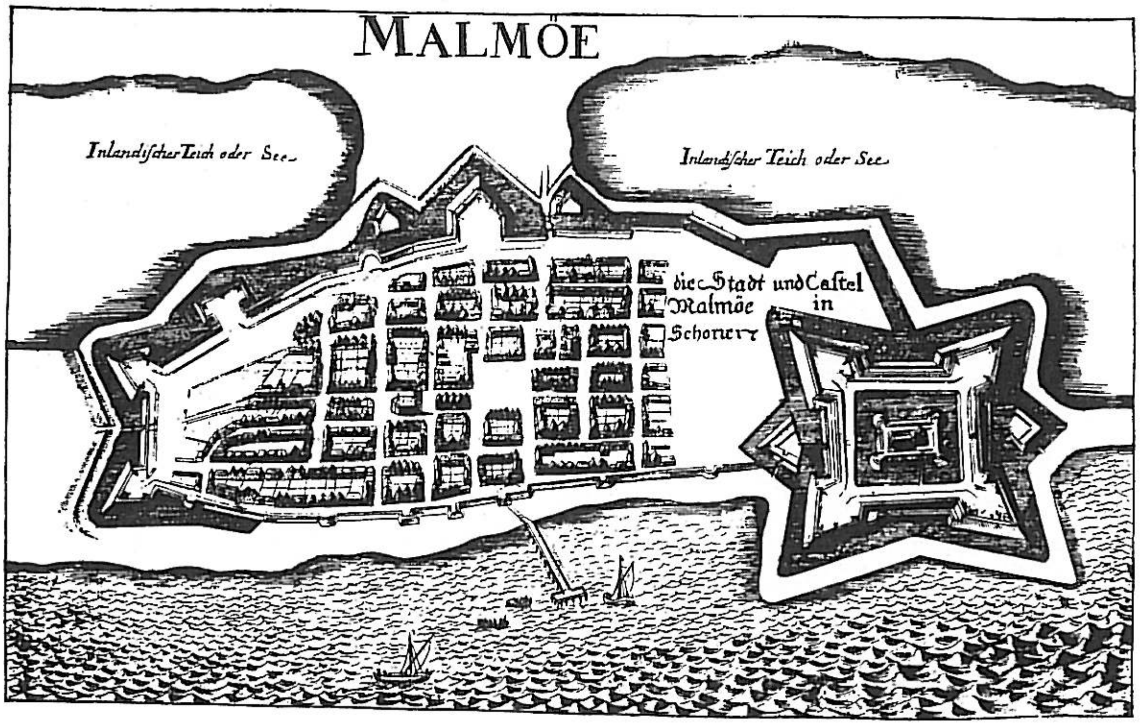
The Malmö Study
A calibration of traffic conflict techniques



A study organised by ICTCT –
the International Committee on Traffic Conflict Techniques

Leidschendam, 1984
Institute for Road Safety Research SWOV, The Netherlands

THE MALMÖ STUDY: A CALIBRATION OF TRAFFIC CONFLICT TECHNIQUES

A study organised by ICTCT - the International Committee on Traffic Conflict Techniques

R-84-12
Leidschendam, 1984
Institute for Road Safety Research SWOV, The Netherlands

This report was prepared by:

G.B. Grayson, TRRL, England

C. Hydén, Lund TH, Sweden

J.H. Kraay, SWOV, The Netherlands

N. Muhlrad, ONSER, France

S. Oppe, SWOV, The Netherlands

Edited by G.B. Grayson

CONTENTS

FOREWORD

In the summer of 1983, traffic safety workers from twelve countries gathered in the ancient city of Malmö in the south of Sweden in order to take part in a unique event. Known to participants simply as the Malmö study, it was the culmination of months and even years of discussion and planning by ICTCT - the International Committee on Traffic Conflict Techniques.

The basic idea of the conflict or near-accident had been around for many years before it was first applied systematically in the traffic safety field in the late 1960s. Since that time conflict techniques have been developed and used in many countries. Although attempts have been made to find an agreed definition of 'conflict', there is still much variation in the terms and procedures used in different countries. The Malmö study was therefore set up to make a detailed comparison of the similarities and differences between the conflict techniques currently in use; in short, to calibrate the techniques against each other.

This report is intended to record the results and the achievements of the Malmö calibration study. Of its five chapters, the first gives a brief background on the development of conflict techniques and explains why a calibration study was needed. Chapter 2 sets out the fieldwork plan, together with short descriptions of the techniques used. Chapter 3 gives the results of the statistical analysis of the data; it is the most important, and accordingly the longest of the chapters. It presents and discusses the results in some detail, but its final section gives a short and non-technical summary of the analysis. Chapter 4 takes a broader view of the results, and reflects the fact that a conflict study consists not only of data collection, but also the interpretation of that data in the diagnosis of safety problems. Finally, Chapter 5 sums up the results and conclusions, and looks to the future.

The Malmö study would not have been possible without the efforts and support of a very large number of individuals and organisations, but two deserve special mention. One is the University of Lund, for organising the fieldwork in Sweden, while the other is the Institute for Road Safety

Research SWOV, for making possible the publication of this report. The ICTCT is grateful to them and to all who helped to make the Malmö study a success.

G.B. Grayson,
Editor

1. BACKGROUND AND DEVELOPMENT

1.1. Introduction

The investigation of traffic safety has traditionally been concerned with the occurrence of traffic accidents and their consequences. However, accidents are rare events and can seldom be systematically observed. The accident potential is still harder to estimate. In many cases the accident frequency is too low to permit reliable estimates, and additional information is needed to get a reasoned statement about the safety of a certain situation.
There are also difficulties with the recording of accidents. Accident data only provide information on recorded accidents, and only a proportion of all accidents is recorded. The recorded accidents cannot claim to be fully representative; some types of accidents may be over-represented, while others may be under-represented. It is sometimes impossible to get sufficiently reliable data from the relatively small number of recorded accidents. The period of time necessary for gathering the amount of accident data required for statistical analyses is often quite lengthy. During such long periods of time the conditions and circumstances under which the accidents occurred have often changed.
The present standard recording systems include only limited information about behavioural aspects of the accident. This situation has frequently been acknowledged at national and international levels.

A logical consequence of these issues is that attempts have been made to supplement accident data with other measures. The first study designed to develop an indicator by which traffic accidents could be predicted, and which could be also employed in order to obtain a better insight into causal factors was that of Perkins and Harris (1967). They identified potential accident situations, which they termed 'traffic conflicts'. Over twenty types of traffic conflicts (or impending accident situations) were defined by them. Essentially these traffic conflicts were identified by the occurrence of evasive actions, such as braking or swerving, which are forced on a driver by an impending accident situation or by a traffic violation.
Since that time traffic conflict techniques have been developed in many

countries and for many purposes, although in most cases the emphasis is still on the diagnosis of safety problems.

1.2. Definitions of 'conflict'

Although attempts have been made to find agreed definitions of conflicts, it is clear that there are still many factors that lead to variation between existing techniques. In practice, the choice of an operational definition of the term 'conflict' seems to depend to a large extent on the objective of the particular investigation. There are already several definitions in use. The most frequently occurring elements of the definitions are the kind of manoeuvre, the proximity of the participants in time and space, the speed of the vehicles and changes in their speeds, the vehicles' direction of motion and changes therein, and the various categories of traffic participants. These elements are usually applied in combination. In most studies conflicts are also classified according to the degree of severity.

There is also a variety of observation methods. With more subjective methods we find terms such as 'sudden behaviour' or 'evasive action' as part of the definition, terms that presuppose a judgement of the observer. Objective methods use terms like 'time to collision' (TTC) or 'post-encroachment time' (PET). Then there is differentiation in the assessment of conflicting behaviour. Terms like 'serious' and 'less-serious' conflicts have been used, referring to the difference in accident potential. The seriousness dimension is usually one dimensional, and only a few techniques use more aspects to define severity.

Based on the investigation by Perkins and Harris (1967), in most American studies the illumination of brake lights and the change of lane are regarded as conflict criteria. However, these criteria do not distinguish clearly minor incidents from severe ones, since the brake lights are seen not only in emergency stops but during normal braking manoeuvres as well.

If one regards the conflict technique as the systematic observation and investigation of risky interactive behaviour, the most important question is which aspects of traffic behaviour are dangerous in which situations.

The usefulness of the conflict analysis technique does not depend on the extent to which accident numbers are correctly predicted, but on whether safety problems can be detected or not. The conflict technique can be seen as a theory about risky interacting traffic behaviour. For this, it is not enough to classify observations as conflicts. One should also specify the seriousness of the conflict with regard to the accident that may result from it. In order to do this, one has to state the relevant cues and the weight these cues give to the seriousness of the conflict. It would be desirable to achieve a greater uniformity in the use of definition elements and conflict types. This would make the comparison of study results easier. However, this problem has not been tackled until now.

1.3. Applications of the conflict method

A study of the literature on the development and application of the conflict method (Kraay, 1983) reveals the following possibilities for application.

The conflict method is based on the assumption that safety decreases as the interaction between road users becomes more and more conflicting. In addition, traffic safety aspects related to conflicts can be analysed as well, thereby revealing the actual causes of safety problems.

As a rule the traffic safety of a given location is characterised by the average annual number of traffic accidents, if possible in relation to a measure of traffic exposure. However, since relatively few accidents happen at any one place, accidents often cannot be used as an exclusive criterion of traffic safety in short-term investigations. It is necessary to look for other data that are related to traffic safety. In other words, it has to be considered whether an accident investigation can be complemented with additional studies in order to reveal the causes and circumstances of likely accidents. The conflict method is an ideal candidate for these purposes.

1.4. Recent developments

In the last 15 years, a number of countries have developed traffic conflict techniques, usually in accordance with their own special circumstances. Based on mutual contacts between research workers (mostly in research groups of the Organisation for Economic Cooperation and Development - OECD) it was found expedient to arrange an international workshop to discuss the various techniques that were being developed at the time.

The first International Traffic Conflict Techniques Workshop was held in Oslo in 1977. Representatives of many of the participating organisations presented papers describing their work on the observation and detection of traffic conflicts (Amundsen and Hydén, 1977).

During this workshop meeting, agreement was reached upon a general definition of a traffic conflict: "a traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged".

The Second International Traffic Conflict Techniques Workshop was held in Paris in 1979, and was intended to update and extend the exchange of information on the use of different techniques (Older and Shippey, 1980).

At the same time the International Committee on Traffic Conflict Techniques (ICTCT) was formally established. One of its main tasks was to decide objectives, plan, design and conduct international studies on the calibration and validation of conflict techniques.

The first international comparative study of conflict techniques was carried out in Rouen in 1979 with the participation of five different observation teams, and reported at the Paris Workshop. The Rouen experiment was considered quite successful by the participants, as they felt their understanding of the different techniques in use had been considerably improved, and that it was both possible and desirable to learn from the experience of other countries. Despite some shortcomings in procedure, the results of the experiment were felt sufficiently encouraging to justify a full-scale calibration study. An organising group within ICTCT was appointed to make the necessary arrangements, and in 1982 the Third International Workshop was convened in Leidschendam (Kraay, 1982). The primary purpose of this workshop was to discuss and finalise a research plan for the joint international study for the calibra tion of traffic conflict techniques that was due to take place in Malmö.

1.5. The Malmö study and its aims

Before starting the fieldwork for the calibration study, a meeting was held in Copenhagen to discuss the latest developments in conflict studies. During this meeting each team participating in the Malmö study described its own technique in detail. The Copenhagen meeting was partly subsidised by the Scientific Affairs Division of NATO, and the proceedings were published as part of the NATO series of ASI publications (Asmussen, 1984).

The primary aim of the study was to make a detailed comparison of the agreement and disagreement between the various observational techniques currently in use, based on data obtained from a field study. Points of specific interest were how well teams agreed in identifying conflicts in a similar way, and to what extent these activities were influenced by location, type of manoeuvre, the road users involved, and so on. A longer term objective was to establish whether data obtained using one technique could be used in a meaningful way by workers with other techniques. This aim is perhaps the more important since it opens up the possibility of greatly extending the data base for research and practice from a national to international level. This is particularly relevant to the question of validation. It would be difficult for any one country to collect all the data needed for a comprehensive approach to the problem of validation, but if it was possible to draw on the results from other countries, then the task would be considerably easier. This was the long-term objective of the Malmö study.

2. DESIGN OF THE STUDY

From a practical point of view, the aims of the study were:

(i) To compare results between the teams with regard to similarities and differences in the scoring of conflicts and the degree of severity associated with the conflicts, and to compare these data with objective measures on the conflicts;

(ii) To require the teams to present a safety diagnosis for each location studied.

2.1. General design

In planning the general design of the study a number of alternatives were considered. One possibility was that films or video recordings might be circulated to the participating teams. Such a technique would be inexpensive due to the limited demand for travelling. Also, the conditions for each team would be very similar, e.g. angle of vision, what could be seen by the observers. However, recording from video or film is not commonly utilised by most of the teams in their normal recording procedure. This disadvantage was considered to be crucial, and the idea of circulating films or video tapes was therefore rejected.

By contrast, recording with manual observers on the ground is regularly used by almost all teams. It was therefore considered as beneficial if the recordings could be carried out in this way, together with simultaneous collection of objective data on traffic situations from video recordings. One of the following strategies could then be chosen:

(i) Recordings in a number of countries at a number of locations in each country;

(ii) Recordings at a number of locations in one country.

The first alternative was considered because then not only variations within a country would be included in the study but also variations between countries. It is obvious that there are large variations between countries regarding general behaviour, interaction between road users, etc. However, the costs for the study would increase considerably if studies were to be made in different countries, and probably not all the teams would be able to participate. This was considered to be more dis-

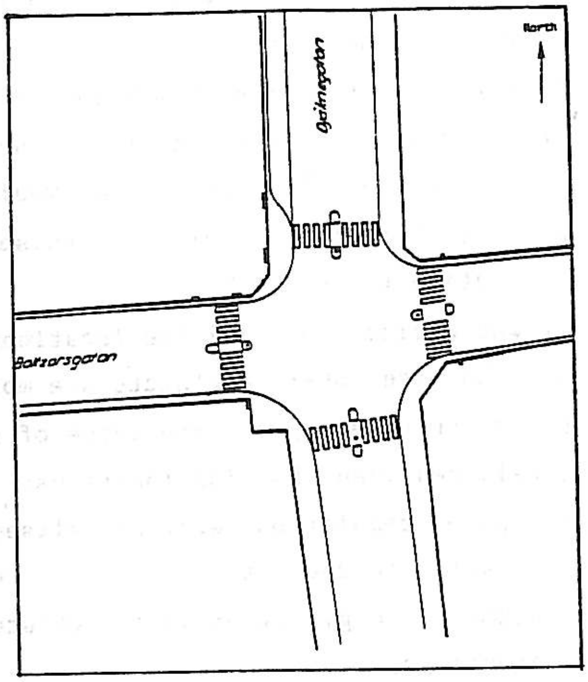
advantageous than to be able to include locations from more than one country in the study.
Sweden was one of the countries that offered to organise the study, and the city of Malmö was chosen for the fieldwork. Malmö is a medium sized city (240,000 inhabitants) that would allow the selection of sites that were as similar as possible to those that might be familiar to the participating teams.
It was decided that all the locations should be intersections as these are the sites where conflicts are most frequent, and because most of the teams would have good experience of such studies. A number of points were considered when choosing locations:

- type of regulation, e.g. signalised or non-signalised;
- location in the city;
- geometry, e.g. street width, existence of median barriers, zebra crossings;
- the practicability of the location with regard to suitable recording sites;
- the time period that the location had remained unchanged.

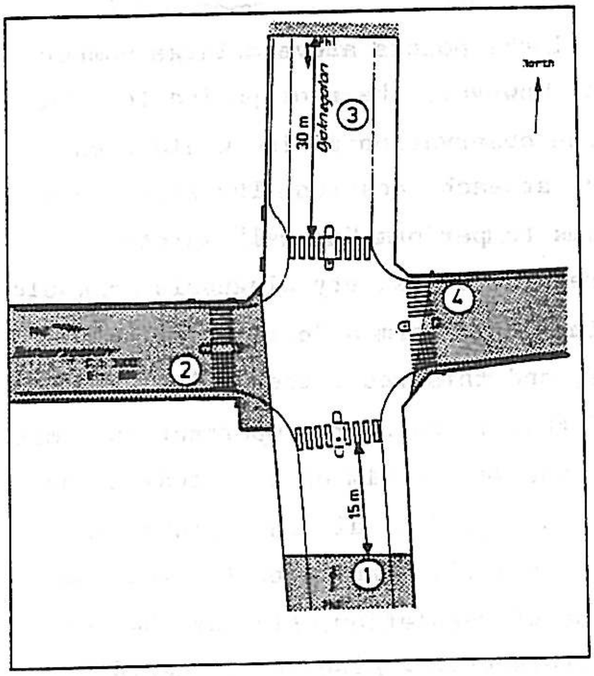
To allow a variation with regard to all the points above a large number of locations would need to be studied. However, the time period for the fieldwork was limited to two weeks. The observation period would then have been limited to less than one day at each location. The second aim of the study - to require all the teams to perform "normal" safety studies and to compare results with regard to a safety diagnosis - would then not be fulfilled. The optimal time to perform a "normal" safety study was considered to be three days, and this would then allow studies at three locations. It was considered that it was more important to limit the number of locations, and to allow the second aim of the study to be fulfilled, than to have a variation with regard to all the points mentioned earlier. Not all the points were equally important. The most important ones were considered to be type of regulation, size of the intersection and the practibility of the intersection. Also it was decided that there should be a mix of road users at each location, and that the intersections should be fairly busy.

Three intersections were finally selected for the study. Details, plans and views of the intersections are given in the pages that follow.

View from Djäknegatan to the north



Sketch of the intersection



Area covered by the video-camera.

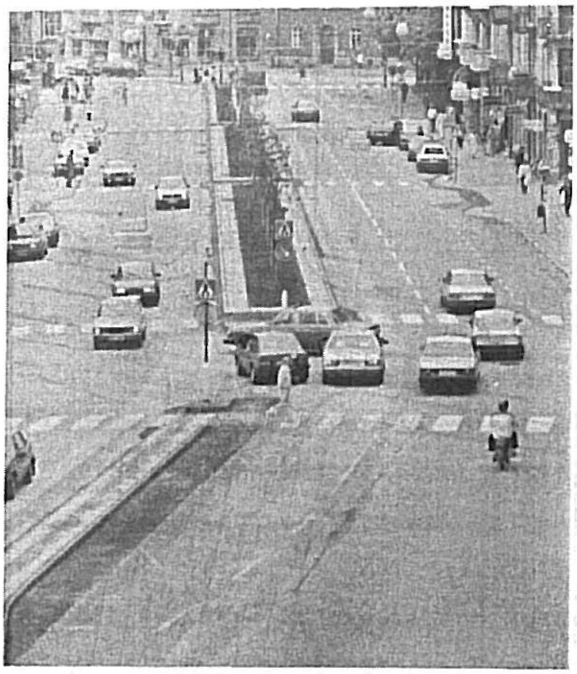
	Djäkne-gatan	Baltzars-gatan
Motor-vehicles (ADT)	13500	3000
Bicycles (ADT)	1000	800
Pedestrians	Busy	Busy
Average speeds on the approaches (Motor-vehicles)	30 km/h	20 km/h

Average annual accident rate: (1978-82)

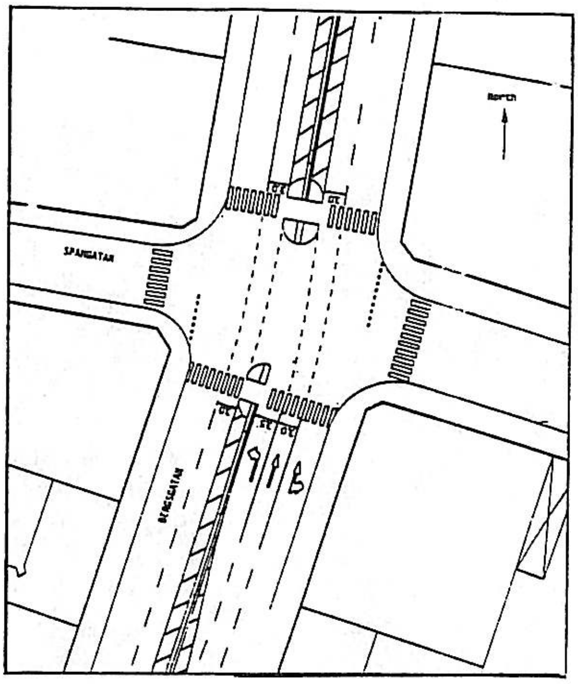
All accidents = 4.2

Injury accidents = 1.0

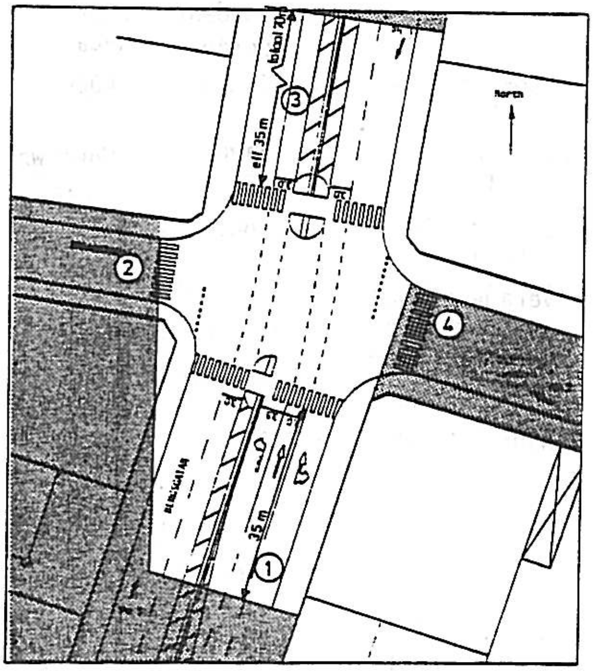
<u>Figure 2.1.a.</u> Intersection 1: Djäknegatan - Baltzarsgatan. Non-signalised. Right-hand priority. Located in the city centre.



View from Bergsgatan to the north-east



Sketch of the intersection



Area covered by the video-camera.

	Bergs-gatan	Spån-gatan
Motor-vehicles (ADT)	11000	4000
Bicycles (ADT)	900	Unknown
Pedestrians	Semi-busy	Semi-busy
Average speeds on the apporaches (Motor-vehicles)	45 km/h	15 km/h

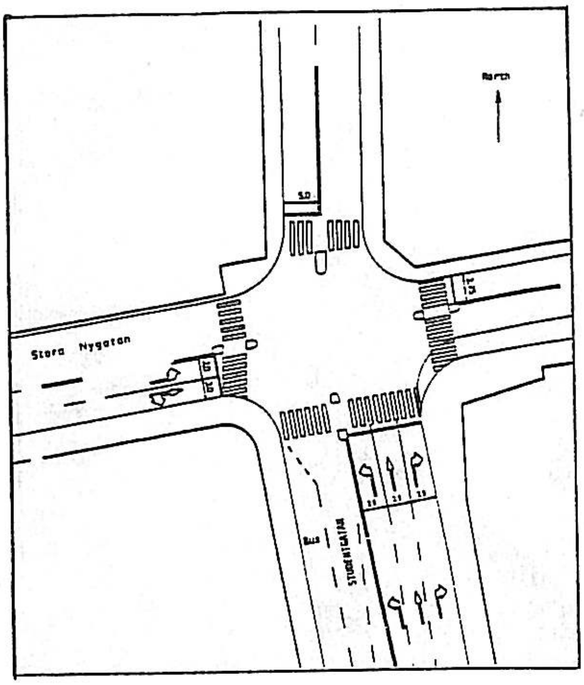
Average annual accident rate: (1978-82)

All accidents = 10.0

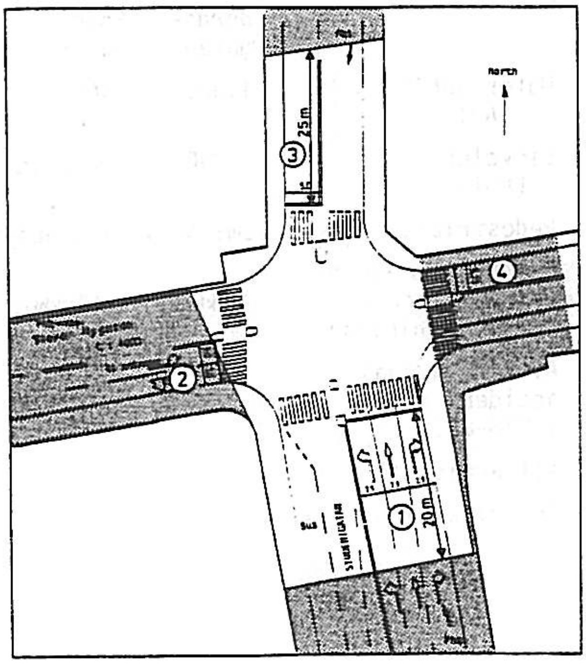
Injury accidents = 2.0

Figure 2.1.b. Intersection 2: Bergsgatan - Spångatan.
Non-signalised. Give way from Spångatan. Semi-centrally located.

View from Studentgatan to the north



Sketch of the intersection



Area covered by the video-camera.

	Student-gatan	Stora Nygatan
Motor-vehicles (ADT)	15000	4000
Bicycles (ADT)	1200	Unknown
Pedestrians (ADT)	Busy	Busy

Average annual accident rate: (1980-82)

All accidents = 7.0

Injury accidents = 1.7

Figure 2.1.c. Intersection 3: Studentgatan - Stora Nygatan. Signalised. Located in the city centre.

Intersection 3 was signalised only two years before the study. This meant that accident data from that intersection was limited. As the study was not a validation study but a calibration one this was not considered to be a great disadvantage.

2.2. Techniques used by the teams

The following conflict teams took part in the study:
Austria: Kuratorium für Verkehrssicherheit (KfV), Vienna
Canada: Transport Canada, Ottawa
Finland: Technical Research Centre (VTT), Espoo
France: Organisme National de Securité Routière (ONSER), Arcueil
Germany: Technical University, Braunschweig
Great Britain: Transport and Road Research Laboratory (TRRL), Crowthorne,
Netherlands: Institute for Perception (IZF-TNO), Soesterberg
Sweden: University of Lund, Lund
USA: Midwest Research Institute, Kansas City

The Danish Council of Road Safety Research carried out accident analysis and behavioural studies.
The Institute for Perception (IZF-TNO), Soesterberg, the Netherlands, collected objective data on speeds, distances, etc. using a video-based technique. They also participated as a conflict team for one of the days.
The Road Safety Centre at Technion, Haifa, Israel, the Department of Civil Engineering, University of Leuven, Belgium, the local road office of Malmö and the Swedish National Roads Administration all had representative present during the study.

The conflict techniques used by the teams are described in the report of the preparatory meeting of the study (Asmussen, 1984) and will therefore not be presented here in detail. However, to provide some assistance in interpreting the results the main features of each technique will be summarised briefly.
Table 2.1 shows the type of definition and the severity scaling used by each of the teams.
As can be seen from the table, Sweden was testing four different scales and France two. Sweden 1 and 2 belong to the original technique, while

	Conflict definition			Severity scaling	
	Estimation of Time to Collision (TTC)	Estimation of Post Encroachment Time (PET)	Interpretation of evasive action	Based on proximity to collision (any type)	Based on proximity to injury accident
Sweden 1 Finland	fixed threshold			X	
Sweden 2	fixed threshold				average speed and type of road user
Sweden 4	threshold function of speed			X	
Canada		fixed threshold	(X)	X	
Great Britain France 2			intensity and result		X
France 1 United States Sweden 3			intensity and result	X	
Germany Austria			intensity and result	X	
Netherlands	calculated minimum value			X	

Table 2.1. Conflict definition and severity scaling used by each of the teams.

Sweden 3 and 4 are new definitions that were tested as part of further development of the technique. Both France 1 and 2 are part of the existing French technique.

Five of the teams had fully developed techniques that are used operationally in their respective countries (Finland, France, Germany, Great Britain and Sweden). One technique could be used on an operational basis (USA) but is not currently in widespread use. Two techniques (Austria and Canada) are in a development stage and are not in operational use. The Netherlands technique is primarily used as a research tool as its costs at present are too high for wider application.

Time to Collision (TTC) is defined as the remaining time to a collision if the speeds and directions of the road users involved did not change. Sweden 1 and Finland estimate TTC via assessments of speeds and distances at the moment of the start of evasive action. Threshold level between serious and non-serious conflicts for both definitions is TTC = 1.5 seconds. Sweden 4 also estimates TTC in the same way, but the threshold level in this case is dependent on the speed of the road user performing the evasive action.
The Netherlands technique calculates TTC curves on the basis of quantitative measurement of the positions of road users in successive frames of the video recordings. From these curves the minimum TTC value is derived. Normally conflicts with a minimum TTC of 1.5 seconds or less are studied.

Post Encroachment Time (PET) can be expressed as the time difference between the moment an "offending" vehicle passes out of the area of potential collision and the moment of arrival at the potential collision point by the "conflicted" vehicle possessing the right of way.
Canada makes a straight estimation of the PET value. The threshold level between serious conflicts and non-serious conflicts is 1.5 seconds. Also Canada scores CE-conflicts (close encounters) which form the highest class of severity.

Interpretation of evasive action is made in different ways by the teams. In the British technique a matrix is used by the field observers, which uses time before possible collision, severity of the evasive action, type of evasive action and proximity in distance when evasive action ends, in order to estimate the severity rate of the conflict. This is normally carried out on site.
The France 2 also uses a risk matrix to transform serious conflicts

(France 1) as predictors of proximity to a collision to risk values predicting the proximity to an injury accident. The main variables in the French matrix are type of conflict and type of road user involved. In the French case the transformation is made after the study.
France 1, the U.S.A., Sweden 3, Germany and Austria are purely subjective estimates of the severity of the conflicts, based on different scales.

All teams do normally use human observers for ground-level observation. The Austrians also apply their technique to the in-car recording of conflicts that the driver of a car is exposed to.

All the teams had to introduce some minor modifications of their techniques for the Malmö study:
- The number of observers varied for some teams compared with what would be the normal procedure at intersections like those studied.
- Observation hours in total were shorter for most teams compared with their normal procedure. This does not affect the comparison of conflicts scored by the different teams, but it makes the safety diagnosis less complete than it would be usually.
- Observed intersections varied with regard to geometry, location and types of road-users for some temas compared with what they normally are used to. The main differences were:
o Canada initially developed the PET concept for specific vehicle - vehicle conflict situations (crossing and turning manoeuvres). Experience of pedestrian and bicycle conflicts was very limited.
o Great Britain is most used to studies at uncontrolled rural or semi-urban road junctions where pedestrians and bicyclists are much less apparent than in the Malmö intersections.
o French observers were not trained to detect bicycle conflicts because bicycles are quite rare under French conditions.
o The U.S. technique is mainly used in semi-urban environments with far fewer pedestrians and bicyclists than was the case in Malmö.

2.3. Procedure

It was decided that all eight teams would perform simultaneously all through the whole calibration study. In the preparatory stage some

objections were raised due to anticipated problems in exposing 15-20 observers at the same location. Observer interaction, it was said, could bias the recordings. It was also stated that so many observers might influence the behaviour of road users. Both objections, however, were discounted. In case of the first one it was concluded that experienced observers were well trained to work by themselves and not be disturbed or influenced by any external stimuli. To emphasise the importance of this question, however, it was stressed before the start of the study that interaction between observers could bias the results and it was therefore important that interaction should be avoided.
With regard to the second point, influence on road user behaviour, it was concluded that the presence of many other pedestrians at the locations and the space available for observers made influence on road user behaviour less obvious. Even though some influence might occur, the main aim of the study, to compare conflicts scored by the different teams, would not be biased by this fact.
The most important reason behind the decision to make all the observations simultaneously was that to split the teams between locations would complicate the comparisons of the results and would increase the duration of the study, something that was most undesirable.

It was considered at an early stage that two weeks was the optimal duration for the study. A longer study might be too expensive for some teams and there was then a risk that these teams could not participate. On the other hand less than two weeks would create too few days for recording. Three days at three locations was considered to be an acceptable distribution of the time available. This allowed one last day for summing up of the study with all observers still present.

With regard to observation hours it was decided that six and a half hours a day was the maximum that the observers could be expected to stand for a two week period under non-domestic conditions. Similar observation periods were planned for the three intersections. Each of the three days per intersection had different hours in order to cover the whole period between 7 a.m. to 8 p.m. Parts of the noon peak and afternoon peak were covered two or three times in order to obtain as many conflicts as possible.

The observation hours were followed exactly by all teams except for the German one, who made 25 minutes of recording and then a five minute interval for each half-hour of observation. The five minutes were used to change the location of the German observers.

Only weekdays were used for observation, again the main reason was to be able to record as many conflicts as possible for the comparisons. The selected days and observation hours were believed to be the most important ones for the safety diagnosis.
Full details of the observation periods are shown in Figure 2.2.

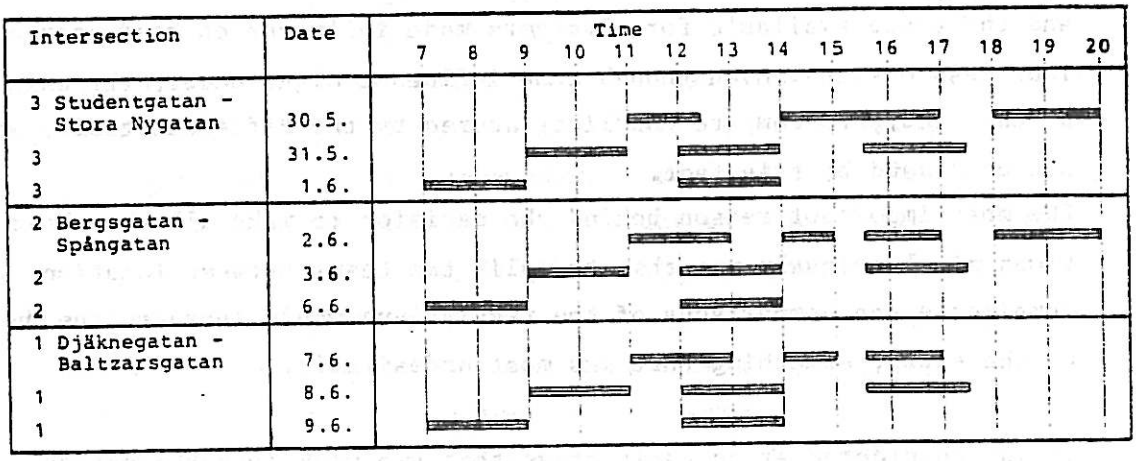


Figure 2.2. Observation periods at the three intersections.

2.4. Recording and labelling

Most teams used two observers on the ground except for Austria who used only one. Canada, Finland and Germany used three observers. All teams except for Canada used observers who were trained in their home countries: Canada trained three observers from Sweden for one week in advance of the study.
The observation area for the observers were defined with regard to what was covered by the video recordings and with regard to what was normally covered by two observers at a study at an intersection.
These areas were as shown in Figures 2.1.a-b-c.

All teams except Austria covered the whole observation area. Austria left out parts of the area, but changed observer location so that the whole area was covered parts of the time.
All observers were located in those approaches to the intersections along which video recordings were made (i.e. the street carrying the heaviest traffic volumes. Most observers were located upstream, 5 to 30 metres away from the intersecting street. Most teams had two observers located diagonally, one at each of the two approaches along the main road. A few teams had observers located downstream.

All observers had to fill in the same data sheet that had been agreed in advance (see Figure 2.3). All teams except for the British one used this data sheet in the field instead of their usual one. The British team transferred their conflicts to the common sheet immediately after each day's fieldwork. The common data sheet contained the data necessary to be able to identify all conflicts (time, road users involved, sketch of the conflict) and severity rate for all conflicts. In order to obtain a high accuracy in the recording of the time for each conflict, all observers were provided with digital watches that were synchronised. Observers were told always to note the time of the conflict to the nearest second.
Space was also available on the data sheet enabling each team to add any information that was considered important or necessary for the scoring of the conflicts or for the safety diagnosis.
After each day of observation the team leader prepared one set of data, as sometimes two members, in accordance with their normal procedure, scored the same conflict on the field. The team leader also checked that all the necessary data was filled in on the data sheet.

Following each day of observation, labelling of the conflicts was undertaken in order to identify each conflict recorded and to give it a unique number. The identification was based on video recordings.
The original plan was that the video recordings would be viewed by representatives of all teams and that on-site observers would not participate in the labelling. This plan had to be abandoned because most teams had only field observers in Malmö and could not allocate a separate representative for the labelling sessions. In the event, the labelling team was usually made up of representatives from four of the teams.

ICTCT - CONFLICTDATASHEET - MALMÖ 1983 | ID-code: ....

Intersection: 1) *Djäknegatan - Baltzarsgatan* .... Date: ....

Team: ☐ A ☐ CAN ☐ D ☐ DK ☐ F ☐ GB ☐ NL ☐ S ☐ SF ☐ USA

Observer: ....

Precise time of conflict

....

hour - min - sec

Road-users involved:

1) ___ ....

2) ___ ....

3) ___ ....

4) ___ ....

C = Car
T = Taxi
P = Pedestrian
B = Bicyclist
M = Moped
Mc = Motorbike
Bus= Public transport
L = Lorry
O = Others

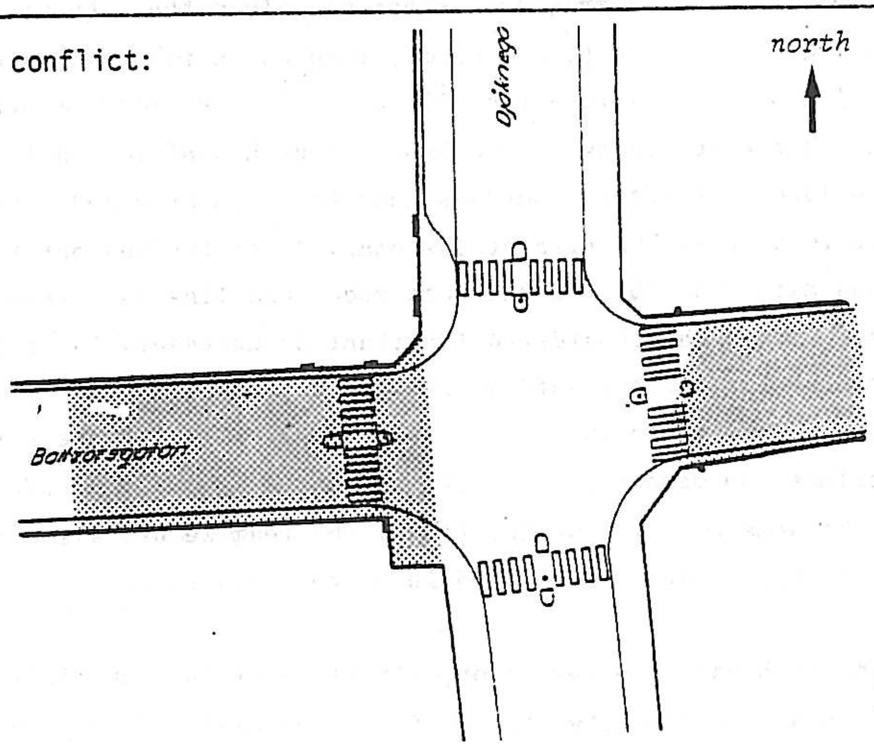
Please note for,

Cars: colour,typ

P/B/M/Mc: age,se

Severity rating: ___ ___ ___ ....

Sketch of conflict:



Not-covered area by video-recording

Please note: Trajectories, number or reference of road-users, particular movem breaking, stopping, skidding, falling etcetera.

Additional data/Comments (to be defined by each team)

Figure 2.3. The common data sheet used in the Malmö study.

After each day of observation the teams collected their data sheets together and sorted them out according to the time of conflict and deleted duplicates. The resulting eight piles of data sheets were then collected and sorted into one pile for each observation period according to the times marked on the data sheets.
The actual labelling then started by identifying the first conflict in the pile from the video tape. The manoeuvre and the types of road users involved were checked in order to ascertain that the data sheet or sheets corresponded to the incident on the video tape. After this the label, i.e. the conflict number, was marked on the data sheets. At the same time a separate list was compiled with the following information: the conflict number, exact time of the conflict, conflict type (road users involved), manoeuvres and the number of teams scoring this incident as a conflict. This procedure was then continued until the last conflict had been identified. The labelling team also checked that all necessary information was written on the data sheets, especially the severity rating.

The greatest difficulties were caused by incorrect times marked on the data sheets, with errors of one and even more minutes being found occasionally. Fortunately, most of these large errors were due to wrongly synchronised watches and were thus systematic. Once the error was identified it could be corrected quite easily for all other data sheets filled in by that observer. Usually, however, the times marked on the data sheets were very accurate, within 10 seconds of the time identified from the video tape.

2.5. Collection of objective data

The objective quantification of road user behaviour was carried out by means of recordings on video followed by a quantitative analysis. For detailed information about the method itself, see Van der Horst (1982, 1983), but a brief outline follows here.

At the three intersections studied two types of black and white recordings were made: one continuously by a normal speed video recorder (Umatic system) (type A), and one by a timelapse video recorder (VHS system) with a reduction factor of 8 (6.25 fields/s) (type B). In the intervals between successive observation periods, the timelapse recorder was running with a speed of 3.125 fields/s (type C). At two intersections, a second camera was used for an insert of one approach to obtain more detail. In principle, the timelapse recordings are not necessary in the analysis process, but were made in case an important traffic conflict occurred during the change of a Umatic cassette (each hour for about 30 s). Afterwards, one slight collision between two cars appeared to have occurred during an intervening period at the third location.

In total, the recordings were made on 48 one-hour Umatic tapes and 7 three-hour VHS tapes. In advance of the study optimal camera positions were selected by the Swedish team. At each intersection the initial installation of the equipment took about one hour, but for the subsequent days filming could be started after about ten minutes. Measurements of distances between 10 reference points on the street were taken after the last observation period of each first recording day, taking between 30 and 60 minutes.

The procedure of the quantitative analysis of video recordings has been described earlier (Van der Horst, 1982). In brief, the analysis consists of selecting positions of some points of a vehicle on successive video stills by positioning electronic cross-hairs. By transformation rules, based on at least four reference points, x and y positions of the video plane can be translated into positions on the plane of the street. Four samples per second appeared to be a reasonable compromise between accuracy and duration of analysis. On average, the plotting of four points of 132 traffic situations with a duration of 16 s took about 36 hours, i.e. 3.7 situations/hour, or an analysis time of 60 times the original duration of the scene. The time for preselecting relevant scenes and describing the characteristics of the road users involved had to be added (about 15 hours, including composing a demonstration tape).

The major purpose of the quantitative analysis was to produce an objective description of conflict situations for the comparison of severity ratings between the teams. This description was in terms of initial speed, maximum acceleration, minimum distance, time to collision, post encroachment time etc, as measured quantitatively from video. The measures used are shown in Table 2.2.

- Road user 1 : Road user with right of way. In car-following situations the first one.
- Road user 2 : Other road user involved in interaction.
- V1 : Initial speed of road user 1 (m/s), as measured in the beginning of the quantitative analysis of the interaction.
- V2 : Idem for road user 2.
- A1 : Maximum acceleration road user 1 (m/s^2) preceding or during the interaction (mean value during one second around the peak).
- A2 : Idem for road user 2.
- MDIS : Minimum distance between road users (m), as measured between two nearest points of both road users before, during or after the interaction.
- TTC : Minimum time to collision value (s). For the used time to collision concept, see a.o. Van der Horst (1982). TTC = 9.9 s means no collision course.
- DTTC : Distance between road users (m) at the moment the minimum time to collision value occurs (99.9 means no TTC value available).
- PET : Post encroachment time (s) after the definition of Cooper (1983). 9.9 no realistic PET value could be computed, mainly because one of the road users had stopped.

Table 2.2. Explanation of parameters, used for description of conflicts.

2.6. Summary

In summary, the following procedure was used in the Malmö study:
- All eight teams that worked with human observers recorded conflicts in the field at the same times and at the same intersections.
- The Lund Institute of Technology made video recordings in order to label the conflicts afterwards.
- The Institute for Perception (IZF-TNO) made video recordings in order to collect objective data on a subset of the conflicts recorded by the other teams, and also carried out an analysis on one day's observations.
- Detailed flow counts were made from the Swedish video recordings by Nottingham University (sponsored by TRRL) and by Lund Institute of

Technology. The results were distributed to the teams to be used in their safety diagnoses if desired.
- An accident analysis was carried out by the Danish team and distributed to all the conflict-recording teams.
- The comparative analysis on all conflicts was carried out by SWOV, the Netherlands.
- Evaluation of the objective data was carried out by IZF-TNO, the Netherlands.
- A normal safety analysis based on the data collected was carried out by each team.

All analyses were carried out during the autumn of 1983 and the results were presented and discussed at a meeting in Berlin in December 1983. Some additional analyses of the conflict data were carried out during the spring of 1984 by SWOV and IZF-TNO.

3. ANALYSIS OF THE CONFLICT DATA

3.1. Introduction

Most of the conflict techniques that have been applied to solve practical problems of road safety are based on subjective observation. It is not always clear what precise instructions are used and how training procedures are organised. In many discussions, especially on an international level, much confusion can arise from this lack of agreement among the various observation procedures and scoring rules. Important questions regarding problems such as reliability, validity, applicability, comparability and generalisation of the techniques, cannot properly be discussed by researchers without a more explicit understanding of the work of others in this practical field.
To solve these problems, ICTCT planned the calibration study in order to create a "universe of discourse" consisting of situations recorded on video that have been scored by teams using the various existing conflict techniques. For a subset of these situations, objective data are present together with background data such as traffic volumes, speeds and accident histories.

From these data it is possible to get an insight into what observers are really doing when they observe traffic, how they select conflicts, what cues are used by them to evaluate the traffic situations in order to select conflicts, and how they arrive at judgements about severity.
This detailed information about their actual observations can be compared with the instructions and can also be used to compare the observations with those of other teams. In such a way it is possible to obtain a detailed comparison of the similarities and differences between teams.
From a more theoretical point of view, it is important to have this detailed information in order to discuss the problems mentioned earlier, especially the problem of validity. In order to progress in this field it is of great importance to have a clear understanding of the scoring rules of observers and the relation of these with the objective aspects of the situation. Only then it is possible to evaluate techniques with regard to traffic safety and to discuss the validity of conflict techniques.

The purpose of this chapter is to provide detailed information on three aspects:

- what is observed by the conflict teams,
- what are the similarities and dissimilarities between these teams with regard to conflict selection and severity rating,
- how are the observations and scores related to objective aspects of the traffic situations that have been observed.

The chapter is organised in six sections. Following this introduction, section 3.2 describes briefly the analytical techniques used in the study. Section 3.3 gives a general overview of the conflict data collected by the various teams. Section 3.4 analyses the subjective scores and assesses the similarities among the eight teams, while section 3.5 investigates the relations between subjective scores and objective data. Finally, section 3.6 gives a short and non-technical summary of the chapter as a whole.

3.2. Analytical techniques

The analytical techniques will not be described in detail, since they have already been reported in the proceedings of the Leidschendam meeting (Kraay, 1982).

For section 3.3, the general analysis, we used the statistical programme WPM (Weighted Poisson Models), a programme for log-linear analysis with an option for weighting cell frequencies with a constant. The weighting option is used in order to correct for unequal observation periods. A description of this programme, which was developed at SWOV, is given by De Leeuw & Oppe (1976).

For section 3.4, the homogeneity analysis, we used PRINCALS. This is a programme for principal components analysis of categorical data, developed by the Department of Data Theory at Leyden State University (Gifi, 1981). This programme handles two fundamental questions at the same time. First, are the scores of the teams homogeneous, i.e. do the teams score the severity of conflicts in a similar way, and is there a common severity dimension? Second, what scales are used by the different teams, and how must we rescale the data of each team in order to compare the individual severity scales? PRINCALS takes categorised variables as a starting point for analysis instead of metric variables (here the categorisation used by a team defines a variable).

The objects (here: conflicts) are completely represented by their profile, i.e. the array of category numbers representing the categories the object is classified in by the various teams. Complete homogeneity between teams can be defined as having identical scores within the profile for each object; objects may differ in scores, but the scores they get from each team must be identical (all variation in scores is between objects and there is no variation within object profiles).

In practice this is only true to some extent; PRINCALS finds the solution that maximises the homogeneity.

Technically speaking, PRINCALS maximises the between-within profile variance. In order to find this optimal solution, PRINCALS replaces the category numbers by category scores. The only restriction for PRINCALS is that, if some category number of a team is replaced by a category score, it must be replaced by that category score in all profiles that have the same category number for that team.

It is proved that this maximal solution is found by Principal Components Analysis on the categories of the variables, instead of on the variables themselves as in classical PCA. There, the substitution of category numbers by category scores is further restricted. If one makes a plot of the category numbers against the category scores, all values are on a straight line. This is the linear (or metric) restriction. In ordinal PCA, the function need not be linear, but must at least be either not-decreasing or not-increasing. With PRINCALS the scores may scatter in the plot (nominal level of measurement). In PRINCALS it is possible to add ordinal or metric restrictions to each of the variables in the analysis. Furthermore we can ask for one and the same quantification in all dimensions. In our case, we assume that each individual team uses one severity dimension. We initially assume a nominal scale for all teams, and after the analysis we can check whether or not order assumptions hold between categories as expected.

In PRINCALS not only the categories get scores, but also the objects. There is a direct relation between object scores and category scores. Once the object scores are known, the category scores can be computed as the mean object score for all objects in that category. On the other hand the object scores are equal to the mean category score of that object, except for a constant. The only further restriction used is a normalisation such that the object scores have a mean equal to zero and variance equal to one. For a mathematical description of the problem and its solution, refer to Gifi (1981). For a simple example of the analysis refer to Oppe (1982). For a comparison of different models of analysis for this example, refer to the Proceedings of the Workshop on Data Analysis in Brussels (in press).

For section 3.5, the comparison of subjective scores and objective measures, we use a programme for non-metric canonical correlation analysis, called CANALS. This is a programme with the same scaling options as PRINCALS. Therefore it is possible to relate nominal variables, such as type of conflict, to the conflict scores, together with metric variables such as TTC's or speeds. Furthermore, it is possible to investigate non-linear relations with CANALS, e.g. by analysing metric variables as if they are ordinal.
The so called "canonical axis" represents some aspect of the group, but not the group as such. The correlation between the canonical axis of the group of subjective scores and that of the group of objective measures, called the canonical correlation,is maximised. If we need a two-dimensional solution, we get a second set of canonical axes that are orthogonal to, and therefore independent of, the previous ones. The canonical correlation between these axes is the highest one under the condition of orthogonality.
We can represent the variables by vectors of unit length in a multi-dimensional space, spanned by the objects. The angles between the vectors correspond to the correlations between the variables; the smaller the angle, the higher the correlation. If we have a two dimensional analysis with only two variables in the second set, then the canonical axes of the second set are vectors in the plane through the two variables. We can project the "explanatory variables" in the first set onto this plane. If a projection is small then this variable does not have much in common with the variables in the second set. Long projections do, however, because they result from small angles of vectors with the plane they are projected onto and therefore represent high correlations.
We will use such a graphical representation in our description of the

results. With CANALS this plot is informative, because the representation is different for different scalings of the variables. The plot then represents the relations between the variables in the first set with those in the second set, after optimal scaling.
In fact, almost all relevant information of a two-dimensional solution can be represented in the plot through the canonical axes of the second set. The canonical axes of the first set project onto the axes of the second set with length equal to the canonical correlation. They are at least as long as the projection on the corresponding axis of any original variable in the first set. We do not refer to the regression weights because it is generally known that these are rather ambiguous. Only the scaling of the variables cannot be represented in the plot. In order to interpret the plot, we always have to relate the findings from it to the scale values of the variables.

3.3. General analysis

3.3.1. The data set

Eight teams, from Austria, Canada, Germany, France, England, Sweden, Finland and the USA, scored the conflicts. France scored each conflict twice, Sweden gave four scores to each conflict. In total 973 conflicts were labelled, of which two were slight accidents. Two of the conflicts were deleted prior to the analysis.
From the 973 conflicts, 111 were scored by at least four teams. Another six received a high severity grade from at least one team. These 117 conflicts were analysed further by IZF-TNO, who computed speed and deceleration of the road users involved, together with minimal distance, distance at minimal TTC, the minimal TTC value and the PET value (see Chapter 2 for definitions). For 14 conflicts it was not possible to compute a TTC score because the road users were not on a collision course. In case of 43 conflicts no PET could be computed, mostly because one of the road users involved came to a full stop.

The total numbers of conflicts at the three intersections were 290, 484 and 199 respectively. For the set of selected conflicts these numbers were 35, 61 and 21 respectively. This chapter is only concerned with the conflicts scored by human observers on the ground. The conflicts scored from video by the Dutch team will be reported elsewhere.

Conflicts were classed by type according to the road users involved, as follows:

(1) Car - car (car includes taxi)
(2) Car - pedestrian
(3) Car - bicycle (bicycle includes moped)
(4) Car - lorry, lorry - lorry (lorry includes bus)
(5) Pedestrian - bicycle
(6) Pedestrian - lorry
(7) Bicycle - bicycle
(8) Bicycle - lorry
(9) Other

Conflicts were also classified by the manoeuvres involved, as follows:

(1) Rear-end conflicts
(2) Weave or merge
(3) Right angle
(4) Head on
(5) Left turn
(6) Right angle with turn
(7) U-turn
(8) Double turn
(9) Pedestrian
(10) Other

Full details with diagrams of the manoeuvre types are given in Annex 2 at the end of the report.

3.3.2. <u>The general analysis</u>

Before we come to the analyses already outlined for the particular aims of the calibration study, we will first make some general comparisons between the data from the three intersections that have been observed and from the different teams that were observing. This will be done with regard to the conflict types and the manoeuvres involved. In addition we will look at trends that may be present in the data.

In Table 3.1 we find the scores summed over all teams for each day of observation and each conflict type. The order of observation was Intersection 3, 2 and 1, and each intersection was observed for three days.

DAY OF OBS. INTERSEC.	CONFLICT TYPE C-C (1)	C-P (2)	C-B (3)	C,L-L (4)	P-B (5)	P-L (6)	B-B (7)	B-L (8)	Other (9)	TOT.	OBS. HOUR
30 May	41	10	14	10	1	0	0	1	2	79	6.5
31 May	57	14	13	4	2	2	1	2	1	96	6.0
1 June	8	1	4	8	0	1	1	0	1	24	4.0
INTERS. 3	106	25	31	22	3	3	2	3	4	199	
2 June	118	39	49	10	0	0	5	1	1	223	6.5
3 June	83	26	40	15	1	1	3	0	4	173	6.0
6 June	43	6	24	6	2	0	3	1	3	88	4.0
INTERS. 2	244	71	113	31	3	1	11	2	8	484	
7 June	59	19	25	4	6	1	5	1	1	121	4.5
8 June	45	12	24	9	3	2	3	2	1	101	6.0
9 June	30	9	10	11	1	1	0	2	4	68	4.0
INTERS. 1	134	40	59	24	10	4	8	5	6	290	
TOTAL	484	136	203	77	16	8	21	10	18	973	

Table 3.1. Total number of conflicts divided according to type of conflict, intersection and day of observation.

In general the number of conflicts recorded decreased with the day of observation. This is primarily the case with Intersection 2, with the largest number of conflicts scored in total. The figures are somewhat misleading, because the observation periods are not all of equal length, as can be seen from the last column of Table 3.1. If we correct for this, then the trend is less but still significant (X^2= 25.9, df=2), and is different for the three intersections (X^2= 19.3, df=4). It should be noted, however, that a simple time correction does not allow for the fact that observation periods, and therefore traffic conditions, were not the same for each day. As far as the different conflict types are concerned, Intersection 2 had proportionally more conflicts of types 1, 2 and 3 but fewer of type 4 compared with the other intersections. The differences between intersections cannot be derived from the time effect, because then Intersection 3 should have had most conflicts of types 1, 2 and 3 and Intersection 1 the least. We may therefore conclude that the differences reflect real differences between the intersections and that any learning effect was a function of the specific locations rather than of the general experimental design. In Table 3.2 we find the data with regard to type of manoeuvre (see previous section).

DAY OF OBS. INTERSEC.	MANOEUVRE 1	2	3	4	5	6	7	8	9	10	TOT.
30 May	15	2	1	0	40	0	5	4	11	1	79
31 May	18	8	0	0	43	0	1	8	17	0	96
1 June	2	2	0	0	15	0	0	2	3	0	24
INTERS. 3	35	12	1	0	98	0	6	14	31	1	199
2 June	10	3	122	1	15	20	8	1	39	2	223
3 June	9	5	106	0	9	10	4	2	28	0	173
6 June	7	2	54	0	5	8	5	0	7	0	88
INTERS. 2	26	10	282	1	29	38	17	3	74	2	484
7 June	10	2	41	0	16	22	0	1	27	1	121
8 June	17	2	43	0	8	9	0	2	18	0	101
9 June	5	6	22	0	3	15	1	4	11	1	68
INTERS. 1	32	10	106	0	27	46	1	7	56	2	290
TOTAL	93	32	389	1	154	84	24	24	161	5	973

Table 3.2. Total number of conflicts divided according to manoeuvre, intersection and day of observation.

From this table we see that manoeuvre type 5 (left turn) is more frequent at Intersection 3, while at Intersections 1 and 2 type 3 (right angle) is dominant. These differences are highly significant and are primarily due to the fact that Intersection 3 is signalised and 1 and 2 are not. We also see that the signalisation effect is not apparent for the other manoeuvre types. There seems to be no particular time effect on the scoring of manoeuvre types.

In order to investigate differences between teams, the conflict types recorded by each team are given in Table 3.3.

COUNTRY	CONFLICT TYPE									TOT.
	C-C (1)	C-P (2)	C-B (3)	C,L-L (4)	P-B (5)	P-L (6)	B-B (7)	B-L (8)	Other (9)	
AUSTRIA	78	32	32	19	3	3	6	6	3	182
CANADA	167	17	79	25	4	1	5	1	5	304
GERMANY	115	43	48	21	1	3	3	4	5	243
FRANCE	68	23	32	15	4	5	5	1	3	156
ENGLAND	224	34	66	41	4	3	5	2	8	387
SWEDEN	41	16	18	9	4	2	2	1	0	93
FINLAND	102	35	52	20	8	4	4	1	3	229
USA	101	36	32	18	5	2	5	2	4	205
TOTAL	454	136	203	77	16	8	21	10	18	973

Table 3.3. Table with the different types of conflicts for the eight teams.

Because of the small number of serious conflicts, no distinction has been made in respect of severity, although it will be discussed later in the chapter. We see that there is a large difference between the number of conflicts scored by each team, but it should be noted that for some teams the analysis included very minor conflicts that would not normally be used in safety assessments. England and Canada scored most conflicts, Sweden and France the least. With regard to the types, for the main types of conflicts, those involving one or more cars we found a large interaction between teams and conflict type ($X^2 = 50.22$, df=21). Canada scored proportionally more conflicts of types 1 and 3 and fewer of type 2. England also scored more type 1 conflicts than the other countries. Austria, Germany and the USA scored higher proportions of type 2, but not significantly so.

In a similar manner, there were large interactions between teams and the manoeuvres recorded, with the Canadian and English teams deviating most from the mean proportions, as shown in Table 3.4.

MAN.TYPE	COUNTRY AUS	CAN	GER	FRA	ENG	SWE	FIN	USA	TOT
1	17	7	21	19	29	11	32	17	93
2	10	8	13	10	9	6	8	5	38
3	62	160	106	49	198	31	81	89	389
4	0	0	0	0	0	0	1	0	1
5	29	81	31	19	41	16	32	27	154
6	19	13	17	15	46	2	13	15	84
7	5	5	9	4	12	3	8	4	24
8	2	7	1	6	5	3	6	2	24
9	36	22	45	33	44	21	47	44	161
10	2	1	0	1	3	0	1	2	5
TOTAL	182	304	243	156	387	93	229	205	973

Table 3.4. Conflicts for each country, divided according to type of manoeuvre

Canada had fewer type 1 (rear-end) and type 2 (weave or merge) conflicts, but more of types 3 (right angle and 5 (left turn). Also type 6 (right angle with turn) and especially type 9 (pedestrians) were low. For England, manoeuvres 3 and 6 were recorded more frequently, and 5 and 9 less frequently. There were also many other differences between the teams, primarily concerning the right angle, the left turn, the right angle with turn, and the pedestrian conflicts.

3.3.3. Summary

Summarising the results of section 3.3, we find that the numbers of conflicts scored decreased with time of observation at a particular site and increased at a new location. Therefore it is considered that the learning

effect was more connected with the location than the general experimental situation. However, the evidence for a learning effect is not conclusive as the three days at each site were not strictly comparable.
As was to be expected from the differences at the locations with regard to signalisation, speeds, volumes and flows, substantial differences were found with respect to conflict type and type of manoeuvre.
Differences between the teams were found in the number of conflicts that were scored. England and Canada scored most conflicts, Sweden and France the least. For conflicts with at least one car involved. Canada scored significantly more car-car and car-bike conflicts and fewer car-pedestrian conflicts. England recorded more car-car and proportionally fewer car-bike conflicts. Austria, Germany and the USA scored a higher proportion car-pedestrian conflicts, but this effect was not significant.
With regard to the manoeuvre type, Canada and England deviated most from the mean proportions of conflicts of a specific manoeuvre type. Canada had less rear-end and weave or merge conflicts and more right angle and left turn conflicts.
There are also many other differences between teams, primarily regarding the right angle, the left turn, the right angle with left turn and the pedestrian conflicts.

3.4. The analysis of subjective scores

3.4.1. Introduction

In order to find out to what extent the observation teams agree with each other in their judgement about the conflicts, a homogeneity analysis has been carried out. Before describing the results, it is necessary to discuss some methodological aspects of this analysis.
What do we mean by "agreement" and how do we measure it? Strictly speaking, two teams agree if they score all conflicts in exactly the same way. However, we know that even if two observers are trained to score conflicts in the same specific way, they will probably not succeed in getting exactly the same scores in real life situations, where they first have to decide whether or not some situation is a conflict before it can be scored. Therefore, we will allow some "noise" in the conflict scores. Agreement then means something like: "two observers agree with each other if they do

not disagree too much", whatever this means precisely. However, this definition is not realistic with regard to the calibration experiment. Teams are not trained to score in the same way, and they use different scoring rules. Some teams start from TTC's, other teams from PET values, still other teams use an overall impression of the conflict situation. There seems to be considerable difference in the scoring of conflicts.

In this study we want to investigate whether or not this difference in approach leads to different observations. In general we may say that all teams agree upon the use of a severity scale. The more severe a conflict, the more probable it is that an accident will result. Some teams distinguish further between the probability of a collision and the probability of injuries.

We will now rephrase our statement about agreement as follows: strategies may differ with regard to severity scoring, but once the severity score has been given, then the variation between scores must be within certain limits, because in principle they are all intended to score the same aspect of danger. If there is a very high amount of disagreement, then questions arise with regard to severity rating.

The first possibility is that teams define severity in different ways. In this case we have to check for more than one severity dimension. An extreme situation will be that each team has its own severity dimension and the disagreement does not result from noise alone, but is primarily systematic. Whether the disagreement is systematic or not can only partly be investigated by means of a comparison of the scores. We can investigate to what extent there is agreement and contradiction. In order to be conclusive about partial agreement and systematic individual differences, we must have additional information about the characteristics of the conflicts. This aspect will be discussed in section 3.5. In this section we will restrict ourselves to an assessment of the agreement between the teams by means of a homogeneity analysis on the subjective classifications.

3.4.2. Homogeneity analysis of all conflicts

For this analysis we included all eight teams, but we decided to restrict ourselves to the primary scores and to delete the second score of France and scores two, three and four of Sweden. A first analysis that included these scores resulted in a solution that was dominated by the trivial fact that the Swedish and French scores agreed completely among themselves in scoring or not scoring.

The complete set of conflicts consisted of 973 observations, scored by at least one of the eight teams as a conflict. Table 3.5 shows the marginals for each team.

TEAMS	CATEGORIES 1	2	3	4	NOT SCORED	NOT OBSERVED
AUSTRIA	168	14	0	0	705	86
CANADA	174	94	36	-	668	1
GERMANY	220	22	1	0	618	112
FRANCE	136	18	1	1	817	-
ENGLAND	338	46	3	0	586	-
SWEDEN	62	25	6	-	880	-
FINLAND	169	51	9	-	744	-
USA	161	42	2	-	768	-

Table 3.5. Marginal numbers for the set of all conflicts.

We analysed the matrix of 8 times 973 classifications by means of a PRINCALS analysis. For the first analysis we chose a three-dimensional solution. It turned out that the conflicts 018 and 900 (the accidents) and conflict 675 (uniquely scored by Germany in category 3) had to be excluded from the analysis in order to get a solution that was not trivially comparing these three severe conflicts with all the others. From the second analysis of the 970 remaining conflicts it was concluded that the agreement between teams was not as high as had been expected. The eigenvalues, which can be regarded as mean squared correlations between the team scores and the common dimensions, are .37, .15 and .12 for the three dimensions respectively.

A comparison of these three values shows us that, if there is agreement about severity, then this agreement is most likely one-dimensional. The second eigenvalue is only slightly higher than the third, and the first eigenvalue is substantially higher. In order to see whether or not the indication of one common dimension is true, we have to check the component loadings. These values show the contribution of each team to the three dimensions. From Table 3.6 it follows that all teams except Canada contribute to the first dimension, but only a few to the other two dimensions.

TEAMS	COMPONENT LOADINGS				
	DIM 1	DIM 2	DIM 3	DIM 1	SQ
AUSTRIA	.56	-.22	-.69	.59	.35
CANADA	-.04	.19	.00	.55	.30
GERMANY	.67	-.32	-.35	.70	.49
FRANCE	.56	.47	.07	.66	.43
ENGLAND	.62	-.29	.44	.74	.55
SWEDEN	.65	.29	.03	.58	.33
FINLAND	.73	.14	.14	.77	.60
USA	.56	-.52	.29	.52	.27
MEAN					.42

Table 3.6. Component loadings of three-dimensional and one-dimensional PRINCALS analysis on the set of all conflicts.

This confirms the indication that there is one and only one common dimension.

In order to find the optimal one-dimensional solution, the same analysis was repeated in one dimension. The eigenvalue of this solution was .42.

From column 4 of Table 3.6 it can be seen that if we compare the component loadings on the first dimension of the three-dimensional solution with the component loadings of the one-dimensional solution, the only difference is that all teams (including Canada) score highly on the dimension.

The component loadings represent the correlations of the teams with this

common dimension. From column 5 of Table 3.6 we see that the mean of the squared correlations is equal to the eigenvalue.
In order to see whether or not this common dimension can indeed be interpreted for each team as a severity dimension, we have to check the category scores for each team. These values are given in Table 3.7.

TEAMS	CATEGORIES					
	1	2	3	4	N-SC	N-OBS
AUSTRIA	0.18	4.77	----	----	-0.13	-0.05
CANADA	-0.22	0.13	2.78	----	-0.11	-0.34
GERMANY	0.17	3.56	13.08	----	-0.18	-0.14
FRANCE	0.36	4.01	9.68	-0.53	-0.16	----
ENGLAND	-0.09	1.98	10.55	----	-0.16	----
SWEDEN	0.82	2.71	3.59	----	-0.16	----
FINLAND	-0.01	1.91	6.34	----	-0.21	----
USA	0.07	2.30	3.58	----	-0.15	----

Table 3.7. Category scores for the set of all conflicts from a one-dimensional PRINCALS analysis.

If we ignore for the moment the categories that represent non-scored conflicts, we see that the scores of the categories that are supposed to be more severe are higher for each team with the exception of code 4 for France, based on only one observation. For categories that correspond to the classes of non-scored conflicts we expect lower category scores. This is also true for all teams except one. For Canada it seems that the non-scored conflicts are, in general, more severe than they should be. This may suggest that Canada has a somewhat idiosyncratic rule to decide on the relevance of conflicts, but if they score, they use the same severity dimension as the other teams. We refer to the findings in the three-dimensional solution, where Canada was the only team that did not contribute to the first dimension. We will investigate this finding further when we compare subjective scores with objective information in section 3.4. In order to show the interrelations between the teams these values are given in Table 3.8.

In fact PRINCALS can be regarded as a normal principal components analysis applied to the correlation matrix of Table 3.8. However, the most important aspect of this analysis is that it results in that scaling of the categories that maximises the principal components solution.

TEAMS	AUS	CAN	GER	FRA	ENG	SWE	FIN
CANADA	.26						
GERMANY	.52	.24					
FRANCE	.21	.37	.29				
ENGLAND	.21	.31	.52	.49			
SWEDEN	.22	.30	.23	.38	.30		
FINLAND	.36	.32	.40	.48	.53	.42	
USA	.32	.15	.36	.11	.31	.20	.35

Table 3.8. Correlations between the conflict scores of the teams for the set of all conflicts after substitution of the classifications by the category scores.

3.4.3. Homogeneity analysis of the selected conflicts

For reasons of economy only a small number of conflicts were selected for further investigation. Only those conflicts that were scored by more than three teams or that received a high severity score from at least one team, have been included.
This selection resulted in 111 conflicts that were scored by more than three teams and 7 conflicts with at least one severe score. One of these conflicts (number 569) was scored by only one team. This conflict was not included in the analysis, because it could not be found on video tape. The other conflicts (including the accidents 018 and 900) have been analysed by IZF-TNO.
From now on we will only be concerned with the analysis of the selected conflicts. First we analysed these conflicts in the same way as the complete set. The analysis in three dimensions turned out to be rather unstable. This was primarily due to the small number of conflicts that was left. In Table 3.9 we see that many marginal values are very small.

Therefore we had to combine some classes. The following changes were made:

018 was removed again

023, score 4 for France became 3

519, score 3 for USA became 2

569, score 3 for USA became 2 (removed later)

675, score 3 for Germany became 2

Although this "data massage" will result in the loss of some information about the most serious conflicts, these are in any case too few in number for reliable conclusions to be drawn about them.

TEAMS	CATEGORIES				
	1	2	3	N-SC	N-OBS
AUSTRIA	50	14	0	46	7
CANADA	25	24	21	47	0
GERMANY	67	20	0	26	4
FRANCE	42	16	2	57	-
ENGLAND	59	33	3	22	-
SWEDEN	35	16	6	60	-
FINLAND	34	39	9	35	-
USA	51	29	0	37	-

Table 3.9. Marginal numbers for the set of selected conflicts.

The second solution was not degenerated, but was not very similar to the analysis of the set of all conflicts. The eigenvalues, .33, .18 and .15 respectively, are not very different, but the component loadings are (see Table 3.10). Also the category scores are different and not completely logical. Therefore we reanalysed the data in one dimension. The eigenvalue is .38. The component loadings are given in Table 3.10, column 4.

If we compare these values with the values of the one-dimensional analysis of the complete set of conflicts (column 4 of Table 3.6), we see that the loadings are very much alike. The value for Germany is somewhat lower, and that for USA higher. Also the category scores are very similar

TEAMS	DIM 1	DIM 2	DIM 3	DIM 1	SQ
AUSTRIA	.62	.32	-.26	.59	.35
CANADA	.06	-.47	.03	.56	.31
GERMANY	.64	.12	-.54	.58	.34
FRANCE	-.34	.08	-.32	.62	.39
ENGLAND	.70	-.02	.49	.72	.52
SWEDEN	-.14	-.65	-.42	.50	.25
FINLAND	.70	.16	.40	.72	.52
USA	.61	.27	-.34	.64	.41
MEAN					.38

Table 3.10. Component loadings of three-dimensional and one-dimensional PRINCALS analysis on the set of selected conflicts.

TEAMS	CATEGORIES				
	1	2	3	N-SC	N-OBS
AUSTRIA	-0.57	1.24	----	0.16	0.52
CANADA	-0.28	-0.30	1.19	-0.23	----
GERMANY	-0.47	1.03	----	0.33	0.57
FRANCE	0.01	1.21	2.40	-0.43	----
ENGLAND	-0.49	0.21	3.57	0.53	----
SWEDEN	-0.11	0.78	1.51	-0.29	----
FINLAND	-0.54	0.37	2.08	-0.42	----
USA	-0.72	0.68	----	0.47	----

Table 3.11. Category scores for the set of selected conflicts from a one-dimensional PRINCALS analysis.

in both solutions (see Table 3.11), especially the categories that represent the conflicts.

The category scores for the categories "not-scored" and "not-observed" are more neutral than in the analysis of all conflicts. This seems logical, because these categories are reduced in number. From Table 3.11

we see that for some teams the category "not scored" is now scaled between categories 1 and 2.

Figure 3.1 shows the agreement in object scores for the conflicts that belong to both sets. Here we also find a high correlation (r=.89). From this plot we see that the conflicts that are selected are more severe. The object scores in each set are normalised, such that the mean value is equal to zero and the variance is equal to one. For the values on the abscissa, the object scores of the smaller set, this is obvious because all conflicts of the analysis are represented. Because the selected conflicts are a subset of the complete set, this need not be true for the ordinate values. We notice a shift of the mean ordinate value in the positive direction. From this we may conclude that the selected conflicts are more serious. This confirms the idea behind the selection procedure: if more than three teams scored a conflict, it will probably be more serious.

On the other hand, we can ask which conflicts that are not selected have high object scores (i.e. >0.5) in the analysis of the complete set. We find out that none of these conflicts has a value larger than the median value of the conflicts in the smaller set. Half of the 27 values that are larger than 0.5 are smaller than 0.7, and only one value is larger than 1.5. Therefore we may indeed conclude that the subset consists primarily of serious conflicts. For this reason, however, it is even more interesting to note that the agreement in both analyses is high because this suggests strongly that the scoring rule is stable over the complete severity range.

Figure 3.2 represents the severity scale for the set of selected conflicts. At the left side of the scale we see the categories again, and at the right side the objects. If all categories had been represented we could have reconstructed the object score for conflict 018. It is clear from the various classifications that it would have a very large severity score.

The correlations between the teams are given in Table 3.12. We note that the correlations for Germany and USA are changed a little, which is in accordance with the changes in component loadings.

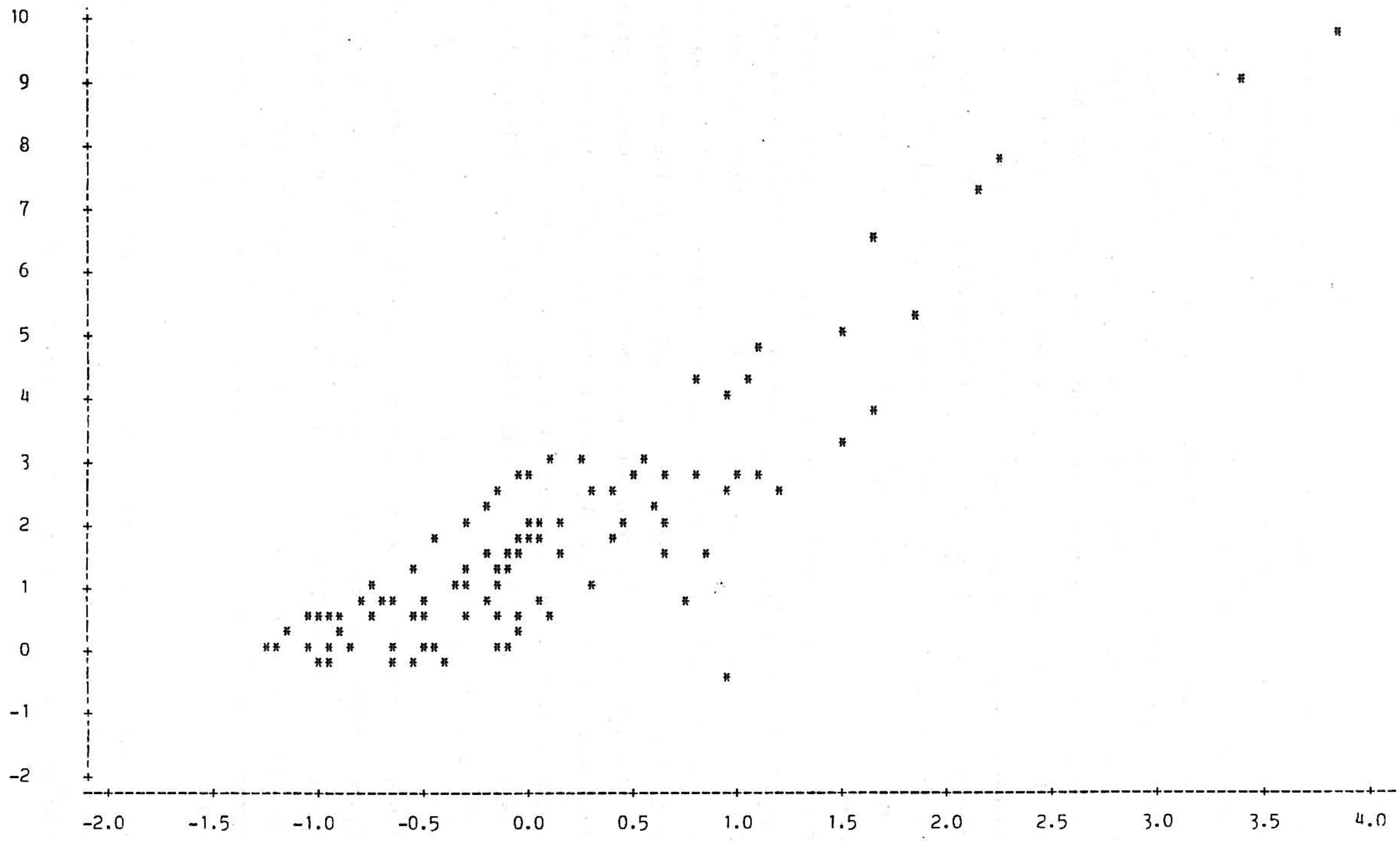


Figure 3.1. Object scores from PRINCALS analysis of all conflicts (ordinate), against those of the analysis of selected conflicts (abscissa).

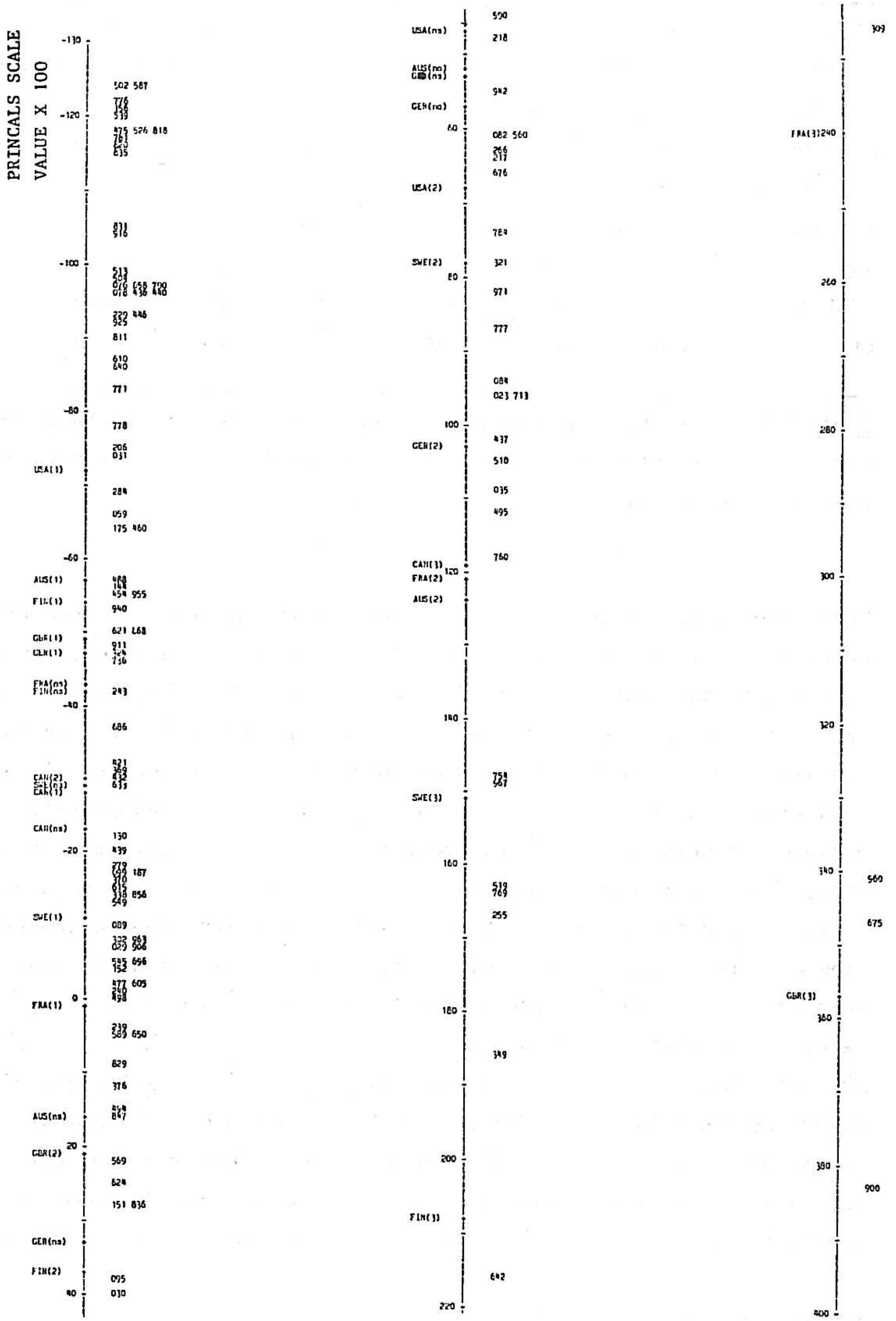


Figure 3.2. Severity scale, with conflicts at the right hand side and categories at the left hand side, from PRINCALS on the set of selected conflicts.

TEAMS	AUS	CAN	GER	FRA	ENG	SWE	FIN
CANADA	.25						
GERMANY	.38	.16					
FRANCE	.16	.38	.25				
ENGLAND	.27	.32	.32	.39			
SWEDEN	.12	.19	.16	.29	.30		
FINLAND	.35	.33	.22	.34	.52	.33	
USA	.40	.18	.44	.26	.31	.20	.35

Table 3.12. Correlations between the conflict scores of the teams for the set of selected conflicts after substitution of the classifications by the category scores.

From Annex 1 we see that there is a reasonable agreement between the conflict scores and the object scores for most teams. However, the not-scored and not-observed conflicts are scattered over the full scale. This indicates that these conflicts are not so much neglected because of non-severity, but primarily because they are not detected as relevant situations. On the other hand we know that this is not completely true, because the selected conflicts are the more serious conflicts, and are on average scored by more teams. Furthermore, we know from the comparison of Table 3.7 and Table 3.11, that this is to some extent due to the restriction of range caused by the selection. Despite this, it seems that the main reason for the low agreement between teams is the difficulty in detection of the relevant situations.
As such, this need not be a serious shortcoming. It only means that observers could be more efficient in selecting situations. However, once the situation is selected, the evaluation is rather accurate. One reason for this difficulty in detection could be the low overall severity of the conflicts, or in other words, the small number of serious conflicts.

3.4.4. Summary

If we summarise the results from the previous analyses we conclude that, although the homogeneity of the scores is not very high, we find one common dimension that can indeed be interpreted as a severity dimension. There is no indication that teams have in common a more complicated severity scoring than the one described by the first principal component. All conflict teams without an exception correlate substantially with this common dimension. Also the reconstructed scales of the teams are in agreement with the expectation and show that we can speak about unanimous severity scaling. The conflicts that were selected for further analysis on the criterion that at least four teams scored them are on the average more severe, and all high severity conflicts of the total analysis are in the selected set.

There is still a considerable amount of difference in scoring. This heterogeneity in scoring derives mainly from the detection of relevant situations, rather than from the evaluation of detected situations. We have already seen that the teams differ considerably with regard to number of conflicts, conflict types and manoeuvre types. However, once a conflict is evaluated, there seems to be a high agreement in the severity rating.

Only a comparison of the scores with the objective characteristics of the conflicts can tell us whether or not the variation in scores that cannot be explained by the common dimension is systematic or random.

3.5. Subjective scores and objective measures

3.5.1. Introduction

The first step in this part of the analysis was to compare the PRINCALS scores with the minimal TTC values as computed by IZF-TNO.

Figure 3.3 gives the plot of these values. We could have added these TTC values to the PRINCALS dataset as if they were the scores of the IZF-TNO-technique, and then reanalysed the data. However, this would not be completely fair because Van der Horst did not select the situations himself as conflicts, and because this solution would be a mixture of subjective and objective information.

The plot shows that there is a relation. The correlation is r=-.46, while for the PRINCALS scores of the analysis of all conflicts r=-.42. If we compare this correlation with the component loadings of the optimally scaled teams, we may conclude that it is rather low. We must, however, realise that the TTC did not get an equal chance and is not scored optimally. Future work is intended on this topic.

From the plot we see that a high severity score always corresponds with a low TTC value. However, the reverse does not hold. Less severe conflicts may also have low TTC values. The minimal TTC seems to be one of the criteria used to evaluate severity, but not the only one. We have to analyse the results further to see what more can be said about the scoring of conflicts by the different teams.

We decided to restrict the number of variables in at least one set. Two approaches have been followed.

In the first approach we examined the extent to which teams used different objective cues. Some teams claim to use the PET value as the only criterion, other teams the minimal TTC, still other teams more complicated evaluations of the conflict type, the manoeuvre and the behavioural aspects. Moreover, it is always possible that the complicated scoring task of observers causes deviations from the scoring rule. The rule at least presumes a certain understanding and interpretation of the situation in order to select conflicts that need to be scored. The observers bear some concept of safety in mind that directs their attention and selection rules. The results from section 3.4 suggest that this task is rather complicated. The first approach resulted in the choice of some CANALS analyses with all conflict teams (including the second French scoring and the other Swedish scorings) in the first set and two objective measures in the second set. For the first analysis we chose the minimal TTC value and the PET value for the second set because many teams refer to one of these measures. From the results of the analyses that will be discussed in para. 3.5.3, we decided to relate the individual team scores also to the conflict type and type of manoeuvre.

The second approach concerns with the explanation of the common severity score from the PRINCALS analysis. Therefore we related the PRINCALS score to all objective measures. In this case CANALS is reduced to multiple non-linear regression analysis. The optimally scaled objective measures in the first set are directly correlated to the optimally scaled PRINCALS score.

In para. 3.5.2 we describe the first, and in para. 3.5.3 the second approach.

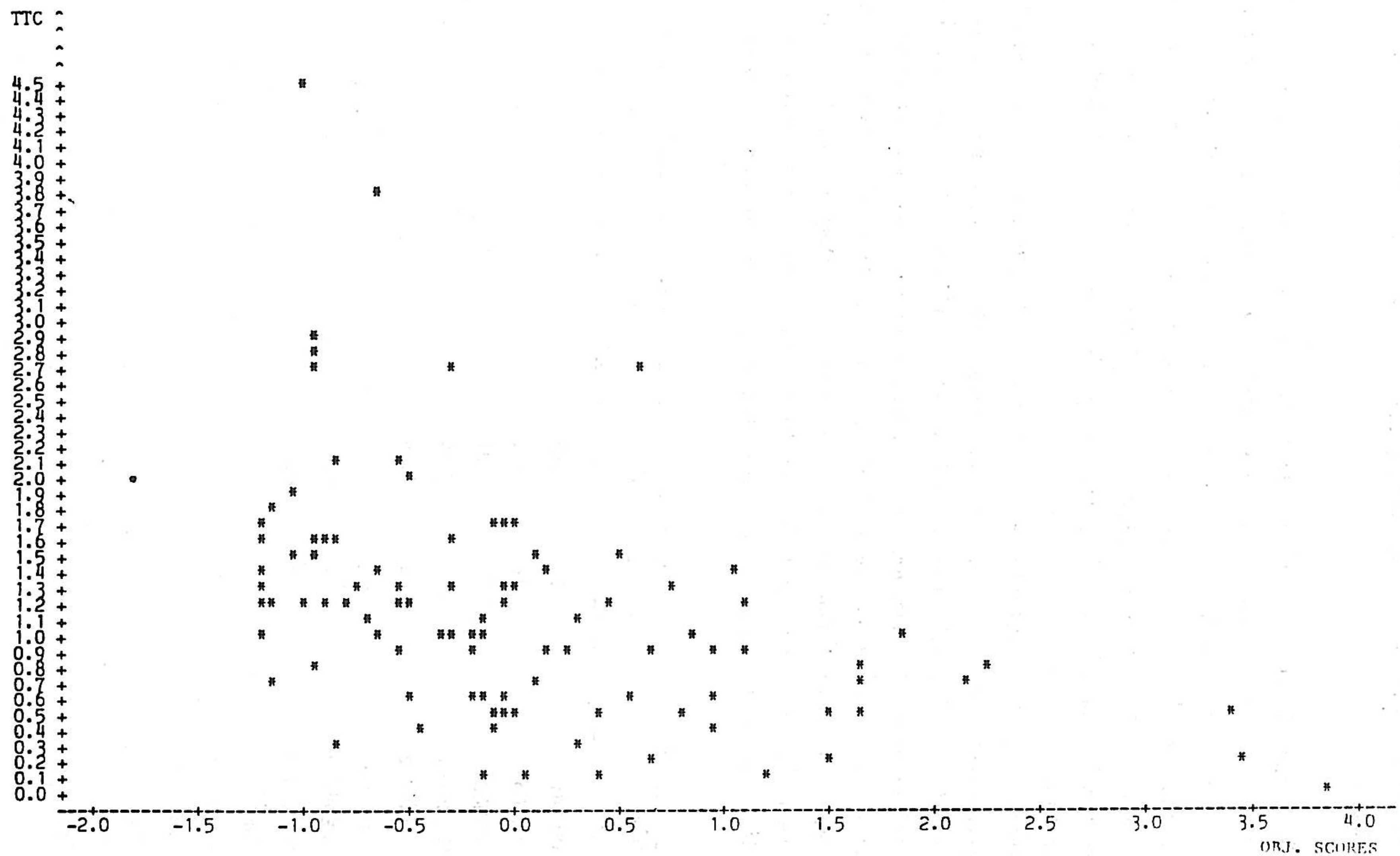


Figure 3.3. Plot of the TTC values against the object scores from the PRINCALS analysis of the set of selected conflicts.

3.5.2. The relation between the original scores and some objective measures

The first CANALS analysis examines the relation between the scores of all teams and the TTC and PET measure in two dimensions. The categories for the teams are restricted to a nominal scale, and the two measures to an ordinal scale. The highest category consists of the missing values for both measures. A first analysis with nominal scale restrictions for both measures showed that these categories did not disturb the ordinal scaling.

The canonical correlations are .78 and .70 respectively, and Table 3.13 gives the correlations of all variables with the two canonical axes of the second set. From these correlations it follows that the first dimension primarily represents the difference between the two measures, and the second dimension the similarity with regard to the scores of the teams. In general the correlations of the variables in the first set with the canonical axes are low. This means that, even after optimal scaling, the team scores do not correlate highly with the TTC and PET values.

TEAMS	DIM 1	DIM 2
AUSTRIA	.22	.19
CANADA	.01	-.13
GERMANY	-.20	-.01
FRANCE 1	.11	-.34
FRANCE 2	.12	-.04
ENGLAND	.09	.17
SWEDEN 1	-.31	-.03
SWEDEN 2	.35	-.08
SWEDEN 3	.22	-.20
SWEDEN 4	.24	-.02
FINLAND	.39	-.09
USA	.29	-.21
TTC	.83	-.56
PET	-.33	-.94

Table 3.13. Correlations of the variables with the canonical axes of the second set.

Figure 3.4 gives a representation of the projections of the optimally scaled variables on the plane through the TTC and PET variables (see section 3.2). In fact these projections represent the correlations of Table 3.13, and the plot is probably easier to interpret than the table. We conclude that the relation between the scores of Canada and the PET values is lower than was expected. High PET values tend to be associated with low conflict scores and not-scored conflicts, but the scaling is irregular, since category 3 is less extreme than category 2. France 1 seems to correlate highly with PET and with TTC as well.
Austria and England correlate primarily with PET and not with TTC. Austria, however, has a positive correlation with PET, while for England we find a negative correlation between the severity scaling and the PET value.
The Finnish and USA teams correlate highest with TTC, but the Swedish scores correlate less with TTC then was expected.
In general we can say that for each team the relation to the common severity score of the PRINCALS analysis is much stronger than the relation to a combination of the minimal TTC and PET values. If there is a relation, then the scaling is not clear in many cases. Furthermore, for no team was the correlation with the best combination higher than .41, and no component loading on the PRINCALS dimension was lower than .50.

From the analysis of the PRINCALS scores and the objective measures to be described in para. 3.5.3, we will see that TTC, minimal distance and conflict type are the most important aspects. In addition to TTC and minimal distance, the conflict type seems to predict a different aspect of the conflict rating. Therefore we compared the scores of all teams with TTC and CT. The canonical correlations were .80 and .72. The category scores for conflict type are -.60, 1.07, .11, -.58, and 3.70.
It was found that the discrimination between the scores is between car-car and car/lorry-lorry conflicts as non-serious on one hand, and car/-lorry-pedestrian and also bicycle-pedestrian conflicts as serious on the other.

From Figure 3.5 it can be seen that the analysis discriminates between the teams of Sweden and France on the one hand and the rest of the teams on the other. The other teams correlate negatively with TTC and positive-

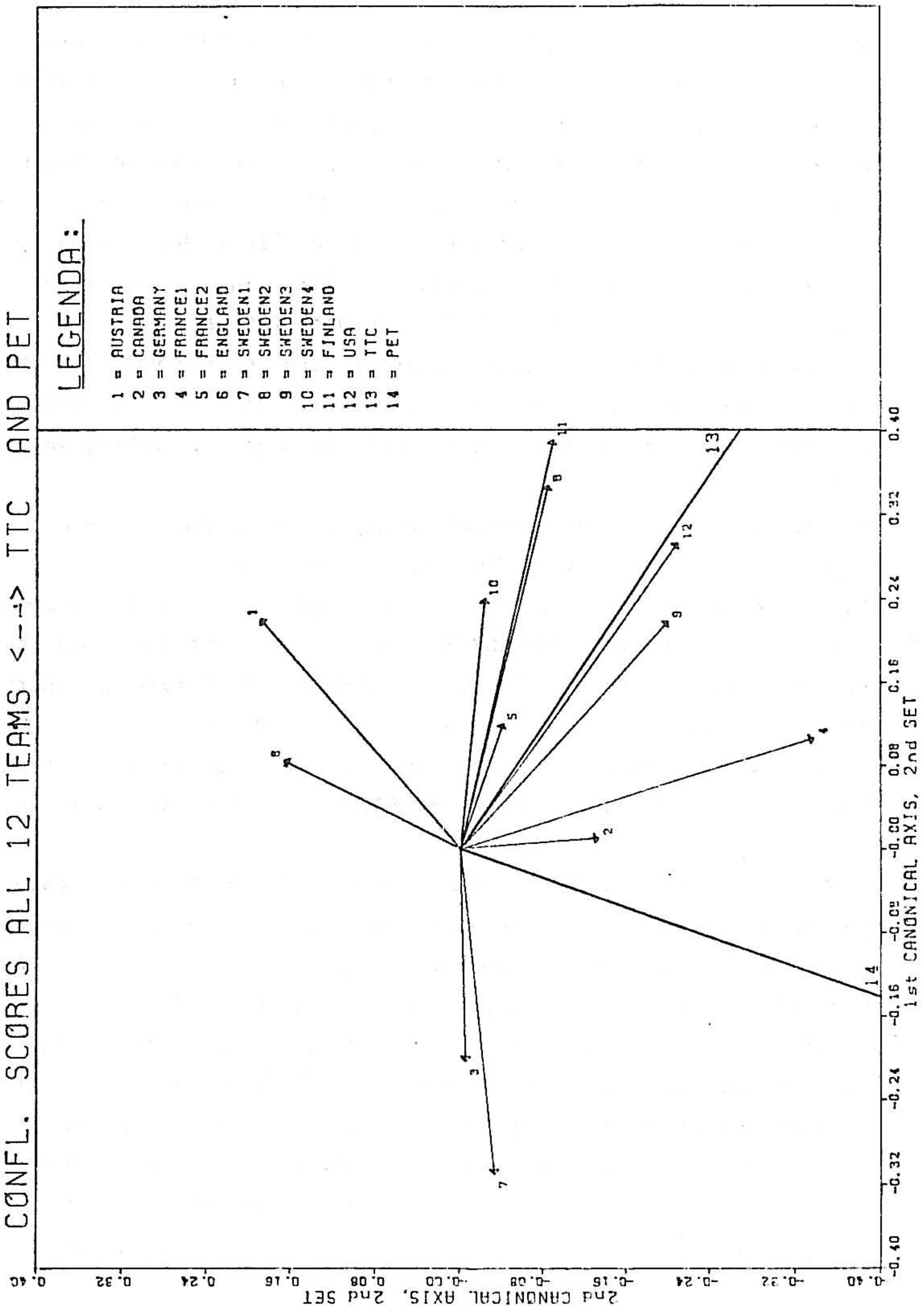


Figure 3.4. CANALS plot of the projections of the optimally scaled teams on the plane through TTC and PET, from the analysis of selected conflicts.

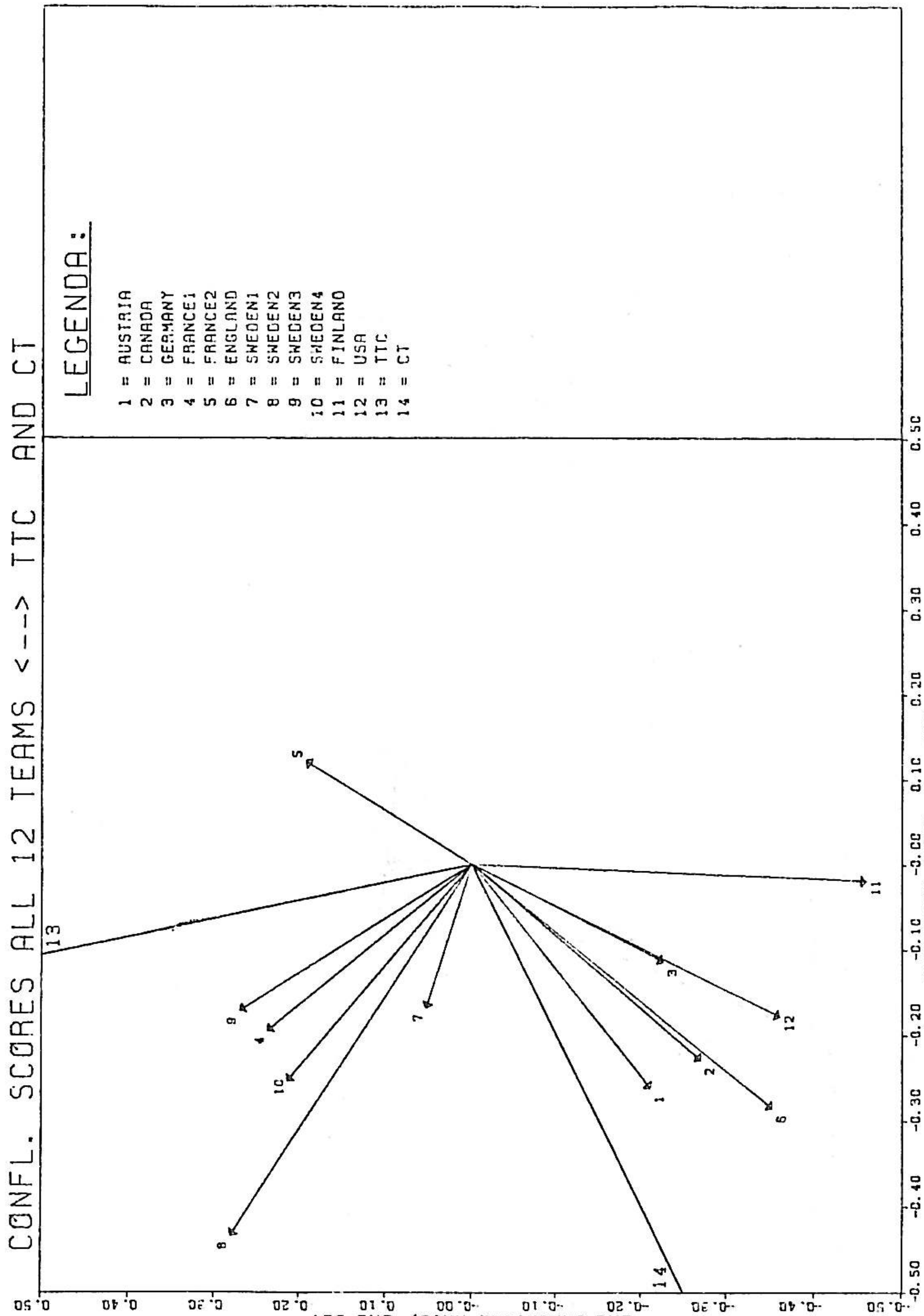


Figure 3.5. CANALS plot of the projections of the optimally scaled teams on the plane through TTC and conflict type, from the analysis of selected conflicts.

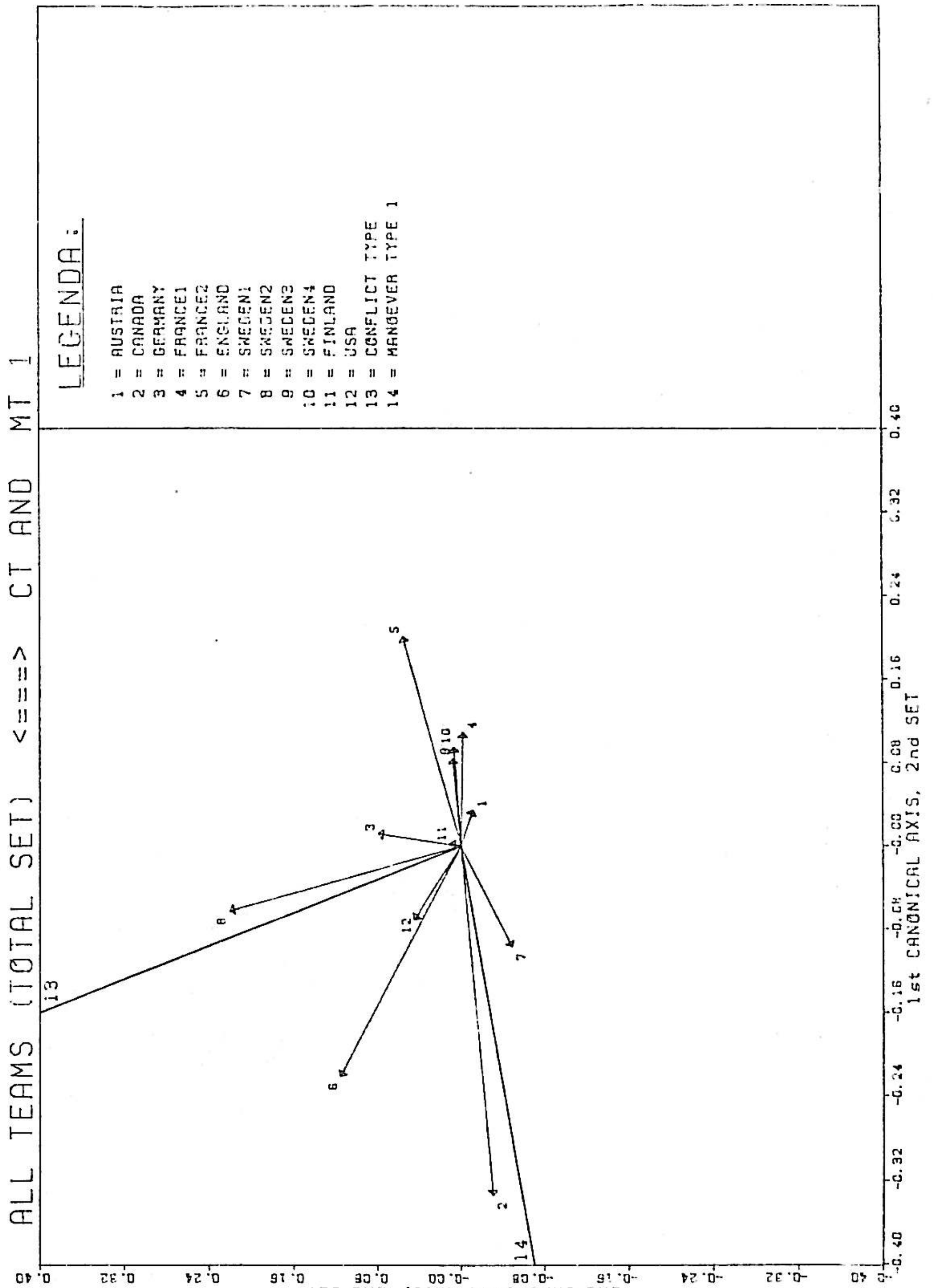


Figure 3.6. CANALS plot of the projections of the optimally scaled teams on the plane through conflict type and manoeuvre type 1, from the analysis of all conflicts.

ly with the discrimination between the CT, as expected, but the teams of France and Sweden correlate positively with TTC. Because of the curvilinear scaling, it is difficult to interpret this result. The distinction between the teams may be somewhat artificial and result from the fact that more than one scaling was used for Sweden and France.

Finally one analysis was carried out on the set of all conflicts. We analysed the team scores as compared to the conflict type and manoeuvre type, and the results are shown in Figure 3.6. The canonical correlations are .54 and .35. These values are rather low, but this is partly due to the stronger scaling restrictions and the larger number of observations. From Figure 3.6 it can be seen that Canada and France 2 discriminate the most with regard to manoeuvre type, Sweden 2 with regard to conflict type, and England with regard to both. Because the correlations are all rather low, we will not go into further detail.

3.5.3. The relation between the PRINCALS scores and the objective measures

The first CANALS analysis of the second approach is a multiple non-linear analysis. The first set consists of the variables conflict type (CT), manoeuvre type (MT), the speeds of the participants (V1 and V2), the decelerations (A1 and A2), the minimal distance (MD) and the distance at minimal TTC (D-TTC). The second set contains the PRINCALS scores.
The scores on all metric variables as well as the PRINCALS scores are classified before the analysis.
The first two variables are nominally scaled, the other variables ordinally. A first analysis showed that we had to recode conflict type because of the small number of observations in classes 6, 7 and 8. Class 6 was combined with class 2, class 7 with 5 and class 8 with 3.
The canonical correlation then was .76. The correlations of the optimally scaled variables with the PRINCALS score are presented in the first column of Table 3.14.
The minimum distance correlates highest with the severity score, and the correlation with conflict type is also high. The correlations with speeds and decelerations are lower than expected, while distance at minimum TTC and manoeuvre have low correlations.

VARIABLES	WITHOUT TTC, PET	TTC, PET INCLUDED
CONFLICT TYPE	.50	.42
MANOEUVRE TYPE	.23	.21
SPEED 1	-.19	-.18
SPEED 2	.02	.01
DECELERATION 1	-.13	-.12
DECELERATION 2	-.23	-.20
MIN. DISTANCE	-.52	-.48
TTC DISTANCE	-.27	-.27
TTC		-.51
PET		-.04
CAN. COR.	.76	.83

Table 3.14. Correlation of variables with canonical axis of second set of CANALS analysis of objective measures vs. PRINCALS scores, without TTC and PET (col.1), and TTC and PET included (col.2).

If we add TTC and PET to the first set of variables, then the canonical correlation increases to .83. The correlations are given in column 2 of Table 3.14. We see that the correlations of the other objective variables do not change much. The correlation for TTC is higher than for the numerical data of Figure 3.3. The scaling is nicely ordinal, shaped somewhat like a logarithmic curve.

The correlation with PET is low, and the ordinal restriction on the scaling is not succesful.

The aspects of the conflicts that correlate highest with the PRINCALS severity score are TTC, minimal distance, and conflict type.

In multiple linear regression analysis there is only a one-dimensional solution. With CANALS, however, it is possible to get a multi-dimensional solution. If we add the dependent variable twice to the second set, then CANALS is free to scale the two variables differently. The two scaled versions then describe two different aspects of the same variable.

Therefore we reanalysed the data with the PRINCALS scores twice in the second set. The first copy was given an ordinal scale restriction, and the second a nominal scale. We further restricted the number of categories of this variable for technical reasons. This was done by combining original categories. We found that the two copies of the PRINCALS scores gave very similar scales. Both scalings, especially the nominal one, turned out to be linear, indicating that the PRINCALS scaling was not an arbitrary one. The canonical correlations are .80 and .68 respectively. Figure 3.7 shows the projections of the explanatory variables on the plane through the two copies of the PRINCALS scores. We see that the two copies are indeed close to each other. If we look at the projections of the end-points of the vectors on the line through one of the copies, then we see that the solution is not very different from the one-dimensional solution.

From the analyses described in para. 3.5.3 we conclude that, as far as teams agree in the scaling of conflicts (and we know from section 3.4 that they do this to a large extent), they use TTC, minimal distance and conflict type as the most important cues. To what extent these cues are used differently for different intersections and manoeuvres is not yet known. For a thorough analysis of these problems more specific data are needed, but in order to get a preliminary indication, some further analyses have been carried out on restricted subsets of conflicts. First we will try to find out what the relation is between the main objective aspects that are important to understand the subjective scoring.
A series of analyses was therefore carried out in order to investigate the extent to which the contribution of the various factors to the explanation of conflict severity is unique, or additional to other factors, what their maximal contribution is, and to what extent they can be replaced by other factors.
The most common way to investigate these effects is by using stepwise regression analysis. Table 3.15 gives an overview of the analyses with regard to conflict type, minimal distance and TTC. The maximal contribution of a variable is found in an analysis with only that variable as a predictor variable. This contribution can be larger than expected (e.g. by means of subtraction) from an analysis with other factors involved. This is due to the fact that the optimal scaling in CANALS need not be

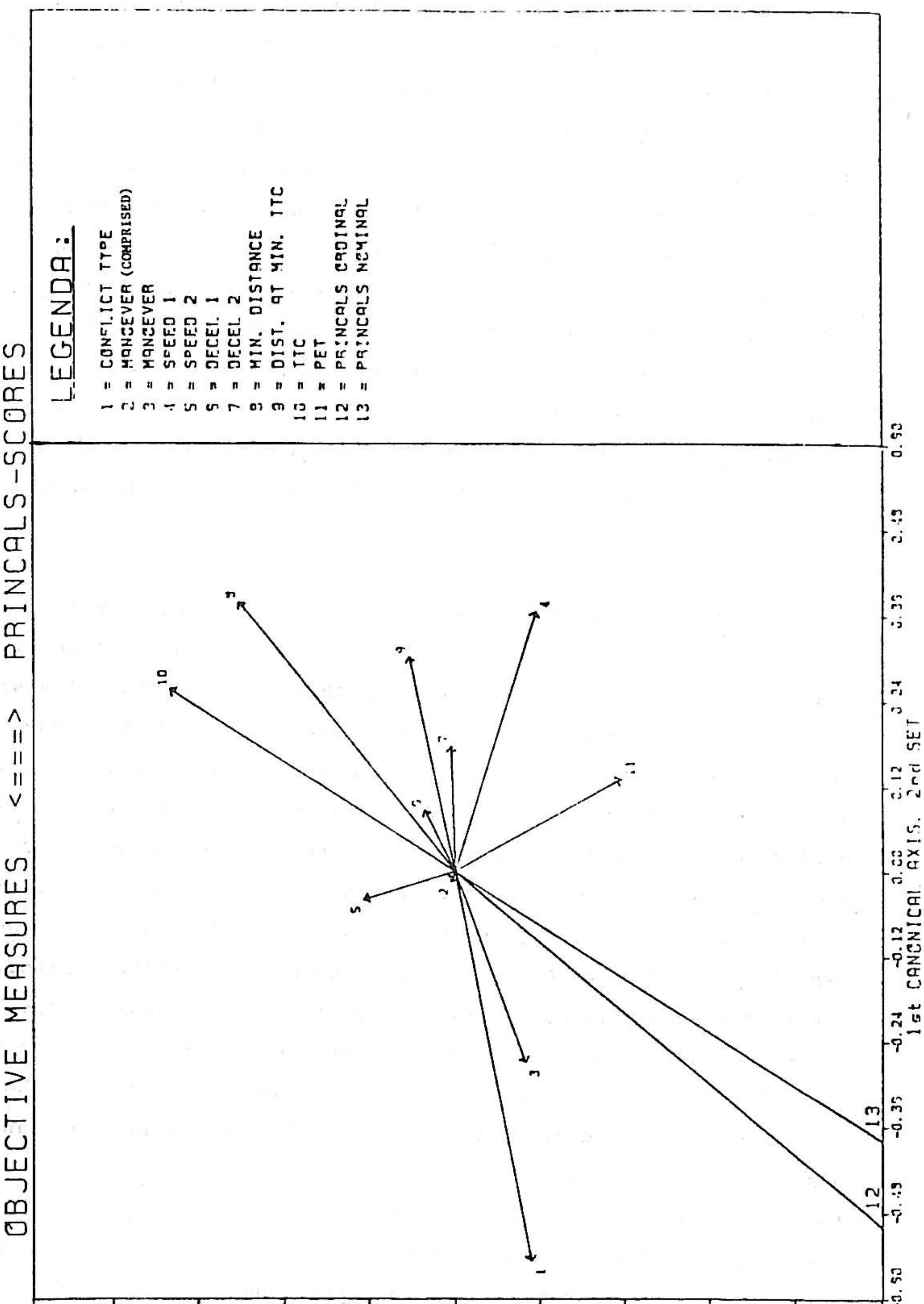


Figure 3.7. CANALS plot of the projections of the optimally scaled objective measures on the plane through the two copies of the PRINCALS scores.

ANALYSIS	CT	MDIS	TTC	CAN. COR.
1	*			.41
2		*		.55
3			*	.60
4	*	*		.60
5	*		*	.67
6		*	*	.63
7	*	*	*	.68

Table 3.15. Multiple correlation coefficients from CANALS all combinations of conflict type, minimal distance and TTC as predictor variables for conflict severity.

the same for the prediction from the variable itself as for the prediction from the group of variables it belongs to in another analysis.
We see from Table 3.15 that TTC is the best predictor, but minimal distance is also fairly good. Conflict type and minimal distance are together as good as TTC is alone. The contribution of conflict type to TTC is larger than that of minimal distance, which seems logical. The contribution of minimal distance to conflict type and TTC is small, but larger to conflict type alone.
This analysis suggests that it is wise to restrict the interpretation of the results of a TTC analysis to a particular type of conflict. Although manoeuvre was not included in this analysis, the same may be true of this aspect also.

3.5.4. The analysis of further restricted sets

It has been shown that conflict type and type of manoeuvre are important cues for observers, for selection as well as severity rating. There are individual differences noticed between teams with respect to this subject. Due to a large variation in type of conflict and manoeuvres, some determining factors or objective cues may disappear in a global analysis. Therefore we analysed more homogeneous subsets of conflicts in order to

check for this. Because of the very small number of objects that remained in these analyses, we only have indications for further research.
The first series of analyses looked at the 55 car-car conflicts. The second series examined a subset of these, comprising 27 right angle conflicts. The analysis of other subsets was not practical because of very small numbers.

In a first analysis the conflict scores of all teams were related to the TTC and PET values. This two-dimensional analysis shows that TTC and PET have rather independent contributions. From Figure 3.8 it can be seen that the relation with TTC is dominated by Sweden, while Canada shows a clear relation with PET in this analysis. The canonical correlations are very high, but this is primarily due to the large number of parameters with regard to the small number of observations. If we run the same analysis in one dimension with only TTC in the second set, then the analysis is still dominated by Sweden 2 and 3, but Sweden 1 does not relate to the description. In general the solution does not differ much from the first dimension of the two dimensional solution as represented in Figure 3.8.

The second group of analyses examines the relation between the common severity score that results from the PRINCALS analysis and the objective measures for the same sets of 55 and 27 conflicts. These analyses are more stable, partly because of the ordinal scaling restrictions for the objective characteristics. These restrictions limit the number of free parameters drastically. Table 3.16 presents the correlations of the objective measures with the severity scores, after their rescaling in the CANALS programme.
In the first column we find the original correlations of the analysis with the total set of 116 conflicts. The second column represents the values for the 54 car-car conflicts and the third column those for the 26 right angle car-car conflicts (one conflict has been removed as an outlier).
From Table 3.16 we see that the solutions for the smaller and more specific sets do not differ much from the total set. The largest change is with speed 1, in this case the speed of the road user on the main road. This speed is reversed for the car-car conflicts, which seems more logi-

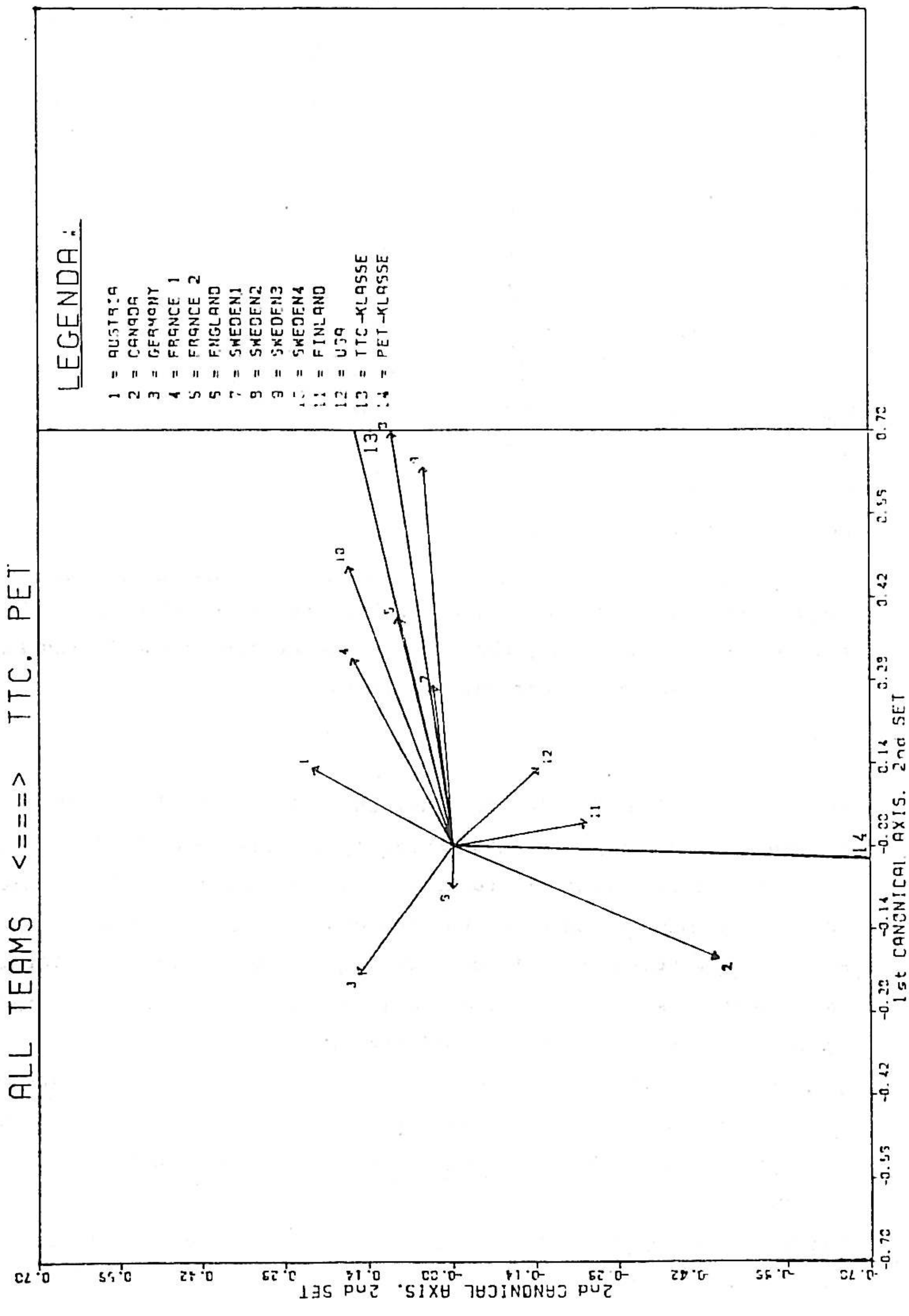


Figure 3.8. CANALS plot of the projections of the optimally scaled teams onto the plane through TTC and PET, from the analysis of 55 car-car conflicts.

VAR	TOTAL (116)	CAR-CAR (54)	RIGHT ANGLE (26)
CT	.42	--	--
MAN	.21	-.01	--
SP1	-.18	.14	.19
SP2	.01	-.27	-.34
DEC1	-.12	-.08	.03
DEC2	-.20	.08	-.43
MDIS	-.48	-.26	-.47
TDIS	-.27	-.31	-.17
TTC	-.51	-.53	-.49
PET	-.04	.25	.32

Table 3.16. Correlations of the objective measures with the PRINCALS scores for the total set, the reduced set of car-car conflicts and the set of the right angle car-car conflicts.

cal. The negative sign in the total analysis may be due to the low speed of pedestrians and bicyclists. Speed 2, in this case the speed for the road user at the secondary road, is now also important, but negative. An explanation may be found in the complexity of the situation, e.g. after a stop or braking for a previous car. This is particularly so for the right angle conflicts. Here also the deceleration of this road user is important, which supports the previous remark.

We may conclude from these analyses that speed and deceleration are important cues, but that the effect depends on the specific situation at hand. It cannot be found unambiguously from the analysis of all conflicts.

Further research must show to what extent the TTC score reflects these aspects. It was shown earlier that minimal distance and TTC are highly correlated, but that minimal distance makes a small unique contribution to the severity score of the subjective observers.

The PET measure becomes more important in the specific analyses, but the effect is opposite to that expected. However, in both analyses the number

of missing PET scores was almost half of the number of observations, while only five conflicts had no TTC value. Nevertheless, it seems that the use of the PET measure is not very practical in situations other than the specific situations for which it was developed.
However, one element of the PET definition could in theory be useful in combination with TTC. In order to use the TTC measure for situations without a collision course, an extension to an area around the vehicle, something like a danger area, seems possible. This area could be defined with regard to speed, deceleration or mass. Perhaps it may be possible in this way to cope better with conflict type, manoeuvre and distance in order to get an improved objective TTC measure.

3.5.5. Summary

If we summarise the findings of paras. 3.5.1 and 3.5.2 we see that TTC is the most important variable. Conflicts with low TTC values are severe conflicts, but not all conflicts that are severe have low TTC values. Other aspects of the conflicts are also important, such as minimal distance, conflict type and, to a lesser degree, type of manoeuvre.
If we analyse the relation between the individual teams and the objective aspects of the conflicts we see that in general the teams relate more with TTC than PET. As far as conflict type is concerned we see that the pedestrian conflicts are regarded as most severe and the conflicts among cars and lorries as least severe. Conflicts between cars or lorries and bicycles are in between. If conflict type and manoeuvre type are related to the team scores we see that the right angle and left turn conflicts are distinguished from the rear-end, and weave or merge manoeuvres and the pedestrian conflicts. Conflicts between cars and lorries among themselves are distinguished again from the conflicts involving pedestrians and bicycles.
Summarising paras. 3.5.3 and 3.5.4, we notice that, if we exclude TTC and PET, then minimal distance seems to be the most important variable for the severity rating, followed by conflict type and type of manoeuvre. If TTC and PET are included, then TTC turns out to be the most important one, while PET does not correlate with severity. Manoeuvre type is less important in this case. The relation between TTC and the severity score seems logarithmic in type. Conflicts between bicycles, mopeds and pedes-

trians among each other are scored as the most severe. Conflicts between cars or lorries and these road users are rather severe, while car-car conflicts are least severe on the average. Right turn conflicts and pedestrian conflicts are severe manoeuvres and the left turn manoeuvres moderate. The right angle manoeuvre is least severe.
As said before TTC is the most important variable, followed by minimal distance, although the latter does not add much to the description of severity by TTC. More is added with conflict type. Minimal distance and conflict type together predict the severity score as well as TTC alone. The fact that speeds and accelerations do not correlate with the severity scoring may be caused by the diversity of the conflicts. Some preliminary analyses with more homogeneous subsets of conflicts showed that in these subsets both aspects are much more important. The subsets, however, are too small for conclusive answers.

3.6. General summary

The analysis of the data collected at Malmö had three main objectives. The first was to establish what had been observed and recorded by the teams in terms of the manoeuvres and the road users involved. In effect this was a general overview of the conflict data. The second objective was to assess the level of agreement among the teams about the selection of events as conflicts and about the severity ratings they were given. The third objective was to relate these subjective scores to objective measures that were recorded for a selected sample of traffic situations.

Different analysis techniques were used for each of the three stages. For the general analysis a log-linear model was used; this incorporated a weighting to correct for unequal observation periods. In the second stage a programme known as PRINCALS was used to establish similarities among the subjective scores from the teams. This is a programme for the principal components analysis of categorical data which has been developed at the Leyden State University. Finally, to compare subjective scores with objective measures a programme for non-metric canonical correlation known as CANALS was used.

Two data sets were employed in the analysis. One was the total set of conflicts observed and recorded by at least one team during the fieldwork period. In the 47½ hours of observation there were 973 incidents that were recorded as conflicts; two of these were minor collisions. Cars were involved in 900 of the conflicts, bicycles in 250, pedestrians in 160, and lorries in 95. The second data set, known as the selected set, was made up of 117 conflicts. These comprised 111 conflicts that had been scored by four or more of the teams, together with a further six that were rated as severe by at least one team. For the selected set of conflicts the objective measures described in Chapter 2 were extracted from video records by the IZF-TNO team.

From the general analysis of the total data set it emerged that the daily numbers of conflicts recorded decreased over time at each of the three locations. This suggests the possibility of a learning effect, but the evidence is not conclusive, as the three days at each site were not strictly comparable in that different observation periods were involved (see Chapter 2).

There were large differences between the numbers of conflicts recorded by the various teams, with the highest scoring team listing over four times as many conflicts as the lowest scoring team. However, it should be noted that some of the teams included in their data numerous minor or less severe conflicts that they would not normally use in an assessment exercise. In addition, a few of the teams were employing techniques that were still in the development stage. Thus the range of scores between established techniques was not as great as it may first appear.
With these provisos, the analysis showed that the English and Canadian teams recorded the most conflicts, and the French and Swedish teams the least. Differences were also found between the teams in the proportions of conflict types (i.e. the road users involved) and of manoeuvres that were recorded. Although statistically significant, these differences were not particularly large, and with a few exceptions the general impression from the analysis was one of agreement in proportions, if not in numbers.

The second stage of the analysis looked at the extent to which the various teams agreed with each other in their judgements about the conflicts they observed and scored. This is a fundamental issue, for a common understanding of the notions of 'agreement' and of 'scoring' is essential to the development of the traffic conflict technique, as well as being central to the calibration exercise itself. A special form of principal components analysis was used for this purpose. The results showed clearly that there was one dominant component to which all teams contributed, and with which they all correlated. There was no indication that the teams used anything other for their scoring than the first component, which can therefore be interpreted as a common severity scale. The analysis was carried out on both the total set and the selected set of conflicts, and it was found that the two solutions were similar. This implies that 'serious' conflicts are not treated differently from other conflicts, they are simply assigned a higher score on the average. The conflicts in the selected set were of above average severity, and all the severe conflicts were present in the selected set.

Although the common dimension in the data is very robust, there are still differences between the scores of the various teams. However, an important finding from the analysis is that the variations in scoring derive mainly from differences in the detection of incidents as conflicts rather than in the evaluations of severity. Once a conflict was scored, then the teams agreed to a large extent on the level of severity. Observers therefore seem to have more difficulty with the detection than with the severity rating of conflicts.

Only a comparison of the scores with the objective characteristics of the conflicts can tell whether the variation in scores that cannot be explained by the common dimension is systematic or is random. This was the purpose of the third stage of the analysis, which was carried out on the selected set of conflicts. If the scores of the teams are related to the objective aspects of the conflict situations then it is seen that the most important variable is TTC (minimum time to collision). Conflicts with low TTC values are severe conflicts, but not all conflicts rated as severe have low TTC values. Other aspects of the conflicts that were found to be important are minimum distance between road users, conflict type and, to a lesser degree, type of manoeuvre. In general, teams relate more to TTC than to PET (post encroachment time). As far as conflict type is concerned, pedestrian conflicts are regarded as the most severe and conflicts among cars and lorries as the least severe.

Further analyses were carried out comparing the objective aspects with the common severity scores, rather than the individual team scores. If TTC and PET are excluded, then minimum distance seems to be the most important variable for severity rating, followed by conflict type and manoeuvre type. If TTC and PET are included, then TTC turns out the most important, and PET does not correlate with severity. The relation between TTC and conflict severity score seems to be logarithmic in type.

In summary, TTC is the most important of the objective measures, followed by minimum distance, although the latter does not add much to the description of severity given by TTC. More is added by conflict type. The fact that speeds and decelerations do not correlate with the severity scoring may be caused by the diversity in the conflict data. Some preliminary analyses with more homogeneous subsets of conflicts showed these variables to be much more important. The subsets however are too small for conclusive answers on this point.

It may be inferred from these results that, even if observers are instructed to use specific cues such as TTC or PET, they will incorporate other aspects of the situation as well. Although severity scaling is linked to objective measures, it also includes a subjective dimension. This results in a common understanding of conflict severity, at least for trained observers.

The analysis of the Malmö data has shown many differences between the teams, but also much agreement. There is a common structure within the data that can justifiably be interpreted as a severity scale, and upon which all teams agree. Differences between teams derive mainly from differences in detection rather than in evaluation. Once conflicts are detected the level of agreement is high, particularly for the more serious conflicts.

4. TEAM RESULTS AND SAFETY DIAGNOSES

The previous chapter has looked in some detail at the results obtained from an analysis of the conflict scores recorded during the fieldwork in Malmö. However, it must be remembered that the detection and recording of incidents in the field is only one part of a conventional conflict study. The other part is the interpretation of that data in order to arrive at an assessment and diagnosis of the safety problems at the sites that have been studied. This is clearly of importance to those who are engaged in the practice rather than the research of traffic safety. Thus the Malmö study was intended to determine both whether the teams agreed in their scoring of conflicts, as analysed in Chapter 3, and whether the teams agreed in their assessments of the safety problems, which will be the concern of this chapter.

The design of the Malmö study required that each team that carried out observations in the field should subsequently provide a report that gave for each intersection an assessment of the safety problems and suggestions for improvements. For the purposes of the calibration study it was the diagnostic aspect that was held to be the more important, since it was unrealistic to expect teams from widely differing backgrounds to be able to prescribe the best solutions to Swedish traffic problems.

Out of the teams that had observers in the field at Malmö, six produced reports that summarised their findings and gave a safety diagnosis of the three intersections. These reports were all written independently, without any team having knowledge of the other reports. The six unpublished reports are:

- Proposals for countermeasures on three intersections in Malmö based upon conflict observations. R. Risser, Road Safety Board, Vienna, Austria
- The national report of the Finnish team. R. Kulmala and K. Salusjarvi, VTT, Espoo, Finland.
- Report of the French team. N. Muhlrad, ONSER, Arceuil, France.
- U.K. results and assessments of Malmö intersections. C. Baguley, TRRL, Crowthorne, Great Britain.
- Swedish report. C. Hydén and S. Almquist, LTH, Lund, Sweden.
- U.S. Summary. W. Baker and J. Migletz, FHA, Washington, U.S.A.

A further three reports were prepared by teams who took part in the Malmö study, but who did not prepare a full safety diagnosis:

- Discussion of the Malmö conflict calibration. P.Cooper, ICBC, Vancouver Canada.
- Report from the Danish team on the results of the behavioural studies performed in Malmö. U. Engel and L. Thomsen, Danish Council of Road Safety Research, Gentofte, Denmark.
- A quantitative analysis of video recordings. A. van der Horst, IZF-TNO, Soesterberg, the Netherlands.

4.1. General comparisons

The team assessments were based on information about the conflicts that had been recorded. However, for comparison purposes the important aspect is not conflict frequencies, but rather proportions. It was shown in Chapter 3 that teams varied considerably in the numbers of conflicts that they recorded, but in this chapter the concern is primarily with how the conflicts were used, and in what proportions. The percentage distribution of conflicts recorded by the teams for the three intersections studied is given in Table 4.1.

COUNTRY	PERCENTAGE OF CONFLICTS RECORDED AT INTERSECTION			TOTAL CONFLICTS RECORDED
	1	2	3	
AUSTRIA	31	50	19	179
CANADA	25	48	27	299
GERMANY	29	54	17	238
FRANCE	27	49	24	153
ENGLAND	27	59	14	379
SWEDEN	37	40	23	93
FINLAND	39	47	14	226
USA	28	51	21	201
ALL RECORDED CONFLICTS	30	50	20	955

Table 4.1. Distribution of recorded conflicts for each intersection.

It should be noted that this table and the others in this chapter make no distinction between the severities of the conflicts recorded; this is a factor that is taken into account in different ways by the various teams, and has been discussed at some length in Chapter 3. However, despite this consideration, which obviously affects the number of conflicts recorded, it can be seen that the overall proportions of conflicts recorded at the three intersections are in good agreement for all the teams.

All teams found the highest number of conflicts at Intersection 2, this intersection was frequently described as having a high accident potential. The other two intersections were not generally considered to have serious safety problems, although the numbers of conflicts collected were high enough to show some deficiences. Intersection 1 was rated as worse than Intersection 3 by all except the Canadian team.
Further evidence of the basic agreement among teams is shown in Table 4.2, which gives the distributions of conflict types recorded by each team summed over all the intersections. The category 'pedestrians' refers to all conflicts involving pedestrians, the 'bicycle' group to all involving bicycles, but excluding bicycle/pedestrian conflicts, and the 'vehicle' category includes all other conflicts.

COUNTRY	PERCENTAGE OF CONFLICT TYPES		
	PEDESTRIAN	BICYCLE	VEHICLE
AUSTRIA	21	25	54
CANADA	7	28	65
GERMANY	20	23	57
FRANCE	21	25	54
ENGLAND	11	19	70
SWEDEN	24	23	53
FINLAND	21	25	54
USA	21	19	60
ALL RECORDED CONFLICTS	17	25	58

Table 4.2. Distribution of conflicts by road users involved.

It can be seen that, with some exceptions, teams agreed well on their allocation of conflicts by road user type.

4.2. Safety diagnoses

All the teams based their safety diagnoses on their analysis of the conflict data, classified according to the type of road user involved and the manoeuvres performed; severity levels, accident prediction, and risk were also taken into account. Most of the teams used descriptive data or comments collected in the field by their observers, and two teams used traffic flows and conflict rates as well as raw figures (Finland and France).
The results are summarised in the following sections.

4.2.1. Diagnosis on Intersection 1, Djäknegatan - Baltzarsgatan

The proportions of road user groups involved in the conflicts recorded by each team are shown in Table 4.3.

COUNTRY	PERCENTAGE OF ALL CONFLICTS RECORDED			TOTAL CONFLICTS
	PEDESTRIAN	BICYCLE	VEHICLE	RECORDED
AUSTRIA	29	21	50	56
CANADA	8	25	67	76
GERMANY	23	23	54	70
FRANCE	24	20	56	41
ENGLAND	14	23	63	102
SWEDEN	29	18	53	34
FINLAND	27	22	51	88
USA	23	14	63	56
ALL RECORDED CONFLICTS	19	25	56	284

Table 4.3. Distribution of conflicts by road users involved at Intersection 1.

All the teams reported that the major problem at this intersection was created by right angle conflicts. This problem was generally experienced by bicycles and motor vehicles, although the Swedish team indicated that it mostly occurred in the latter category.
Pedestrian conflicts were noted as frequent by all teams. Most pedestrian conflicts were reported to occur as a vehicle on a straight path was leaving the intersection. Other types of problems mentioned were conflicts involving a turning movement, and rear-end or weave conflicts; these two categories were considered to be more a sign of operational difficulties than a real indication of danger. The British team also pointed out that cyclists travelling close to the edge of the road were often not noticed immediately by turning drivers.

All the teams attributed the conflict situation to a problem of right-of-way. The solutions suggested were, either to put in traffic signals synchronised with those at the next intersection or to clarify the priority situation with a stop or a yield sign. The traffic signal solution was indicated as likely to reduce pedestrian conflicts as well as right angle ones, provided the phasing is adequate. The British team also suggested the introduction of delineated cycle lanes at the intersection, and the Swedish one the installation of speed humps on all approaches. The French team recommended that a check is made for visibility problems that might account for the variation in conflict rate on different sections of the intersection.

4.2.2. Diagnosis on Intersection 2, Bergsgatan - Spångatan

All teams agreed that the right-angle conflict was the most important problem at this intersection, and that it presented a high accident potential. These conflicts were shown to involve bicycles as well as motor vehicles. The proportions of road user types involved in the conflicts recorded by each team at this intersection are given in Table 4.4.

Bicycles were generally considered to be highly at risk at the intersection. According ot the Swedish and Finnish teams, bicyclists were mostly at risk when travelling on the main road on a straight path. Pedestrian conflicts were not so frequent, but they were still mentioned

COUNTRY	PERCENTAGE OF ALL CONFLICTS RECORDED			TOTAL CONFLICTS
	PEDESTRIAN	BICYCLE	VEHICLE	RECORDED
AUSTRIA	19	27	54	89
CANADA	2	40	58	142
GERMANY	20	23	57	127
FRANCE	17	31	52	75
ENGLAND	9	19	72	225
SWEDEN	16	32	52	38
FINLAND	16	30	54	107
USA	24	23	53	103
ALL RECORDED CONFLICTS	16	26	58	476

Table 4.4. Distribution of conflicts by road users involved at Intersection 2.

as a potential danger by most teams. By contrast, conflicts involving a left turn were shown as frequent but minor by most teams, and similar comments were made on weave and rear-end conflicts. Finally, operational problems linked with U-turns were mentioned by two teams. The situation was unanimously attributed to the priority given, "de facto", to the main approaches of the intersection, the high speeds of motor vehicles, and the width of the central reservation being too small to enable cars to stop in the middle while crossing or turning left. One solution proposed was the installation of traffic signals. All the teams recommended speed reducing measures; a decrease of speeds could be obtained by narrowing the roadway, using the extra space either to widen the central reservation or to introduce a cycle lane or a cycle track.

4.2.3. Diagnosis on Intersection 3, Studentgatan - Stora Nygatan

On this signalised intersection, the main problem turned out to be opposing left turn conflicts. All the teams stated that this was the major hazard for drivers, and two teams indicated that they involved bicycles as well as cars.

The proportions of road user types involved in the recorded conflicts are shown in Table 4.5.

COUNTRY	PERCENTAGE OF ALL CONFLICTS RECORDED			TOTAL CONFLICTS
	PEDESTRIAN	BICYCLE	VEHICLE	RECORDED
AUSTRIA	15	23	62	34
CANADA	16	11	73	81
GERMANY	15	22	63	41
FRANCE	24	19	57	37
ENGLAND	13	12	75	52
SWEDEN	29	14	57	21
FINLAND	19	19	62	31
USA	12	17	71	42
ALL RECORDED CONFLICTS	16	19	65	195

Table 4.5. Distribution of conflicts by road users involved at Intersection 3,

All the teams except one noted a pedestrian problem, with conflicts occurring mainly as a vehicle turned right or left. Other pedestrian conflicts involved crossing against red, and a minor accident of this type was recorded during the observation period.

Less important problems were also noted, including weave or rear-end conflicts, left turn conflicts other than the category already stated, and U-turns. Bicycles were recorded by the Austrian team as being frequently involved in conflicts, especially against vehicles turning right, while for the British team it was left turn conflicts that often involved bicyclists.

All the teams proposed as a countermeasure a minor modification in the timing of the traffic lights; the addition of a special phase for left turning vehicles was sometimes also recommended. Moreover, the Austrian team suggested a bicycle lane across the intersection, the British one a ban on U-turns, the Swedish one a move of the pedestrian crossings nearer to the intersection, and the French one a ban on some turning manoeuvres in the intersection.

4.3. Accident data and some comparisons with conflicts

As mentioned earlier, accident data comprising all police-reported accidents during the five-year period 1978-82 for Intersections 1 and 2 were available, and some relevant accident data for Intersection 3 were also included. Frequencies of those accidents which occurred within the same observational areas used in the conflict studies are given in Tables 4.6, 4.7 and 4.8 and diagrammatic location and vehicle manoeuvres involved are shown in Figures 4.1 to 4.6.

At Intersection 1 the unprotected road users were only involved in 22% of all accidents but 60% of the ones involving personal injury. Pedestrians seem to be most vulnerable as they were involved in all injury accidents with unprotected road users. In almost half of all accidents one of the road users involved was obliged to give way to another (right-angle accidents).
When compared with the safety diagnosis based on the conflict studies (see also section 4.2) it is found that all teams stated that the right-angle conflicts was the major problem at this intersection.
The pedestrian problem was detected by all teams, and considered by some to be as frequent as right-angle ones.

At Intersection 2 the unprotected road users comprised 19% of all accidents but 78% of the injury accidents. 53% of all accidents and 67% of the injury-producing ones were right-angle accidents.
There were nine car-car accidents (one injury) of rear-end type where the first car stopped for a pedestrian.
When compared with the safety diagnoses of the teams it was found that all teams agreed that the right-angle conflict was the main problem at this intersection. Most teams also indicated that these conflicts involved bicyclists as well as motor vehicles. 20% of all accidents of

ROAD USERS INVOLVED	P-C	B-C	C-C	Total
ALL ACCIDENTS				
Number	3	1	14	18
Per cent	17	5	78	100
INJURY ACCIDENTS				
Number	3		2	5
Per cent	60		40	100

Table 4.6. Intersection 1, Djäknegatan - Baltzarsgatan. Distribution of police-reported accidents in the period 1978-1982.

ROAD USERS INVOLVED	P-C	B-M	B-C	M-C	C-C	TOTAL
ALL ACCIDENTS						
Number	1	1	4	3	38	47
Per cent	2	2	9	6	81	100
INJURY ACCIDENTS						
Number	1	1	4	1	2	9
Per cent	11	11	45	11	22	100

Table 4.7. Intersection 2, Bergsgatan - Spångatan. Distribution of police-reported accidents in the period 1978-1982.

ROAD USERS INVOLVED	P-C	B-C	C-C	Total
ALL ACCIDENTS				
Number	1	3	11	16
Per cent	6	19	69	100
INJURY ACCIDENTS				
Number	1	2		4
Per cent	25	50		100

P = Pedestrian, B = Bicycle, C = Car or other motor-vehicle.

Table 4.8. Intersection 3, Studentgatan - Stora Nygatan. Distribution of police-reported accidents in the period October 1980-1982 (after traffic signals were installed).

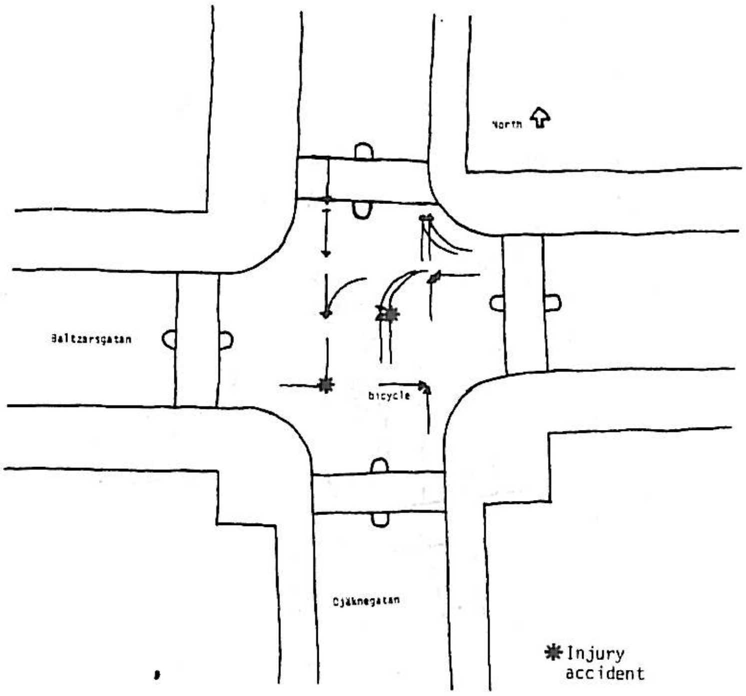


Figure 4.1. Intersection 1, Djäknegatan - Baltzarsgatan. Accidents between vehicles going straight ahead or making a turning manoeuvre.

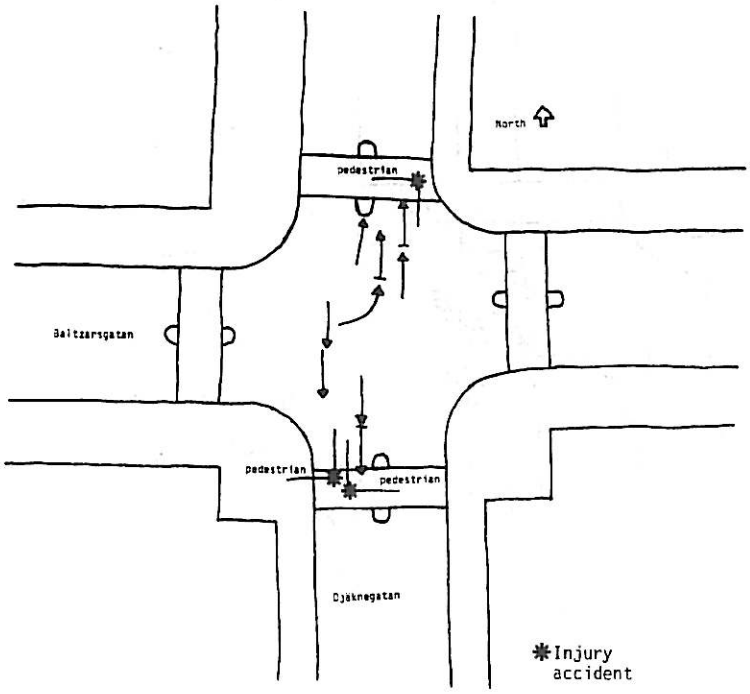


Figure 4.2. Intersection 1, Djäknegatan - Baltzarsgatan. Accidents involving pedestrians directly or indirectly, and single vehicle accidents.

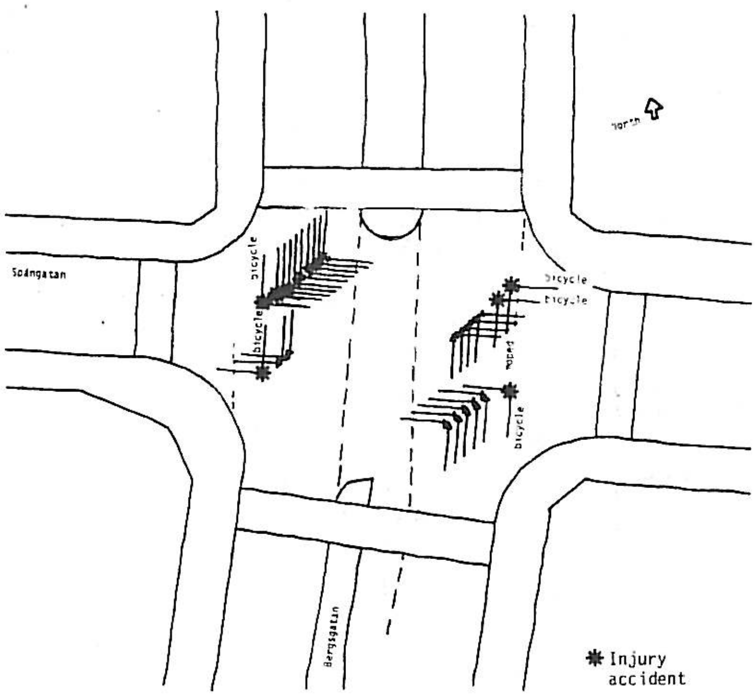


Figure 4.3. Intersection 2, Bergsgatan - Spångatan. Accidents between vehicles going in two directions, where one is obliged to give way to the other and where all vehicles are travalling on a straight path.

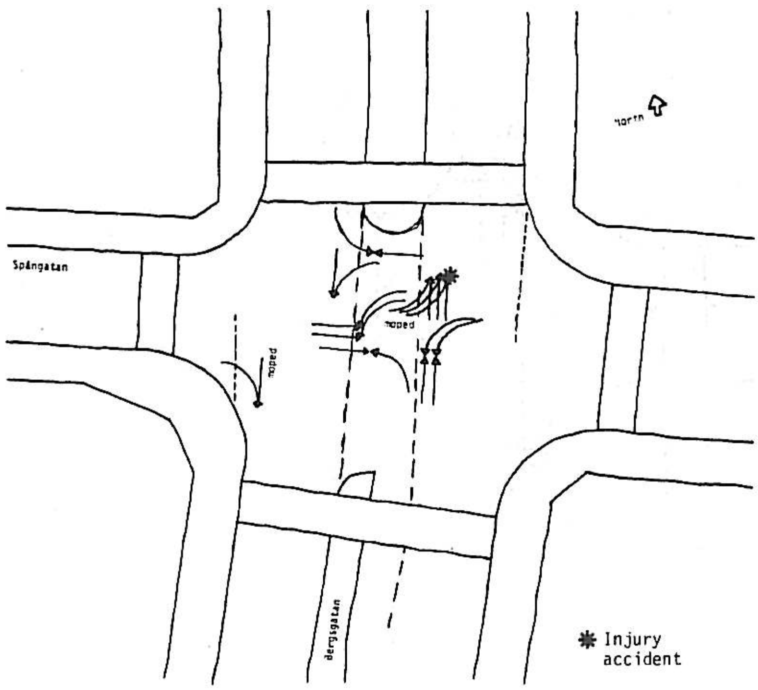


Figure 4.4. Intersection 2, Bergsgatan - Spångatan. Accidents between parties where one vehicle is making a left turn or a right turn.

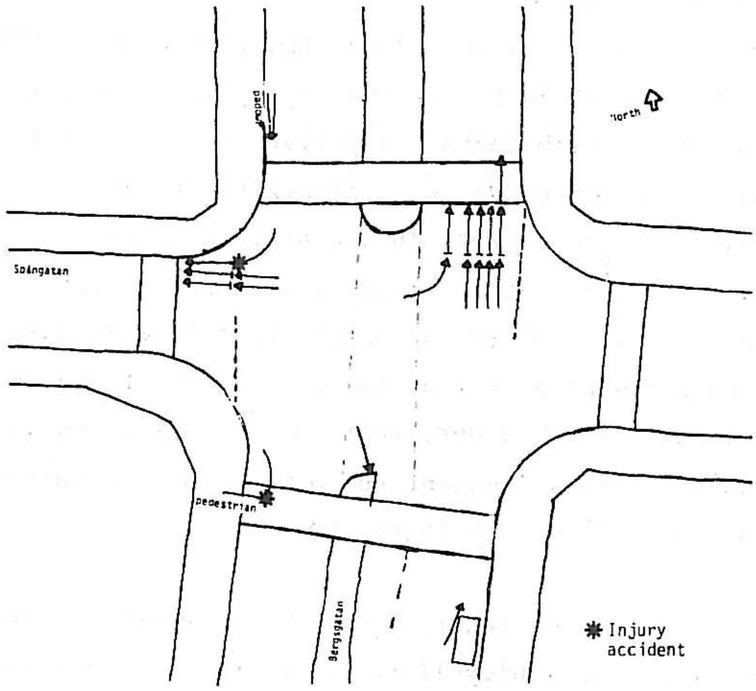


Figure 4.5. Intersection 2, Bergsgatan - Spångatan. Accidents involving pedestrians directly or indirectly and single vehicle accidents.

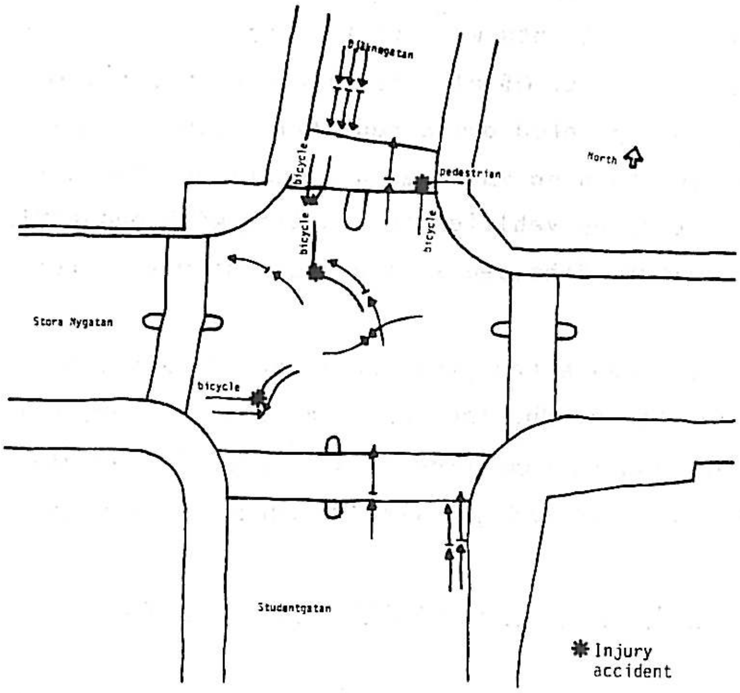


Figure 4.6. Intersection 3, Studentgatan - Stora Nygatan. All accidents.

right-angle type involved a bicyclist while the corresponding percentage for injury accidents was 83%.
Bicyclists were generally considered as highly at risk at this intersection. They also comprised 56% of the injury accidents. According to the Swedish and the Finnish teams, bicyclists were mostly in difficulty when travelling on the main road on a straight path and 3 of the 4 bicycle-car injury accidents were indeed of this type.
Pedestrian conflicts were not as frequent at this intersection but they were still mentioned as a potential danger by all teams apart from Sweden. They comprised 2% of all accidents and 11% of the injury ones. Conflicts involving a left turner, rear-end and weave conflicts were mentioned by many teams as frequent but minor. These comprised about 40% of all accidents and 22% of the injury ones.

At Intersection 3 all three injury accidents after the signalisation involved a pedestrian or a bicyclist. Seven out of 11 car-car accidents were of the rear-end type.
When compared with the safety diagnoses of the teams it was found that all teams considered the main problem to be opposing left-turn conflicts. USA, Sweden and Britain indicated that they involved bicyclists as well. Two out of 4 injury accidents were of the opposing left-turn type and they involved a bicyclist. Of all the 15 accidents, 3 were of this type. All teams except one pointed out a pedestrian problem and one of the three injury accidents also involved a pedestrian. The main problem was considered to be turning vehicles in conflict with pedestrians walking against a green signal (the pedestrian accident was of this type).

In conclusion, it appears that most teams detected the major problems at all three intersections. The accident numbers, however, were too small to make any reliable conclusions from these alone, but provided some indications of safety problems and qualitative comparisons with conflict counts

4.4. Adaptations of the techniques for the Malmö study

Some modification of techniques in order to take part in the Malmö study was inevitable, and this could have detracted from the quality of the results.

The choice of experimental locations was a problem for some teams. For the U.S. one, the fact that two of the intersections were without traffic signals created difficulties, as the American technique has been mostly developed on signalised intersections. For both the U.S. and the Canadian teams, the traffic mix was different from home conditions, and the Canadian technique in particular has not been designed to relate to pedestrian conflicts. For the French team, used to observing mopeds in traffic, the detection of bicycle conflicts was difficult. The British team indicated that their technique had been developed mainly on rural junctions, but adaptation seemed to be easy in this case. The fact that the number of observers was limited to two or three created a problem for the Canadian and the Germans who were used to larger groups in the field.

The universal data sheet was considered difficult to fill in by the Americans and, to some extent, by the French. The limitation of the observation field to the area covered by video was unsatisfactory for the Canadians and the Germans, who stated that most weave or rear-end conflicts were omitted in the data collection. The German team also had to organise the coverage of the observation field in a way that was different from usual practice.

The observation periods were considered longer than usual by the Finns, and also by the Germans who had to take five minute breaks at regular intervals. By contrast, the total observation period on each intersection was shorter than the British and French teams were accustomed to; this was not a problem for calibration, but it could have made the safety diagnoses less reliable.

The risk matrix used by the French team had been tested under French conditions and it should have been adapted to Swedish ones, but this was not possible during the short time available. Risk values obtained from the Malmö data were therefore doubtful, but conflict rating was not affected.

4.5. Adaptations as a result of the Malmö study

One of the immediate effects of the Malmö calibration study was to enable each team to reassess its own technique. While in general teams were satisfied with their own performance, they still found possibilities for improvement. The Americans will revise their severity scaling, and the Swedish and the French their risk matrix. The Finns intend to modify their data sheets into a form resembling that used in Malmö. The Swedes, following the Germans, will also improve the ergonomics of the observing task and introduce short breaks in the observation periods. Austria will give further attention to the threshold between light and serious conflicts, while Canada will improve its conflict definition with a tentative association with PET and TTC. The Netherlands intend to look further into the relation between objective and subjective measures.

Apart from the Canadians, improvements found useful by the different teams will not change the basis of their techniques, which will remain comparable to what they are now, but rather will aim at increasing the reliability and validity of the conflict technique.

5. CONCLUSIONS

The two previous chapters have presented at some length the detailed result of the Malmö study. Chapter 3 concentrated on an analysis of the data collected in Malmö, comparing the scores of the various teams, and relating them in turn to objective data. Chapter 4 then looked at the way in which the teams interpreted the data that they had collected in Malmö. The purpose of this chapter is to summarise all these results, and to assess their implications for the traffic conflict technique both now and in the future. For ease of presentation, these results will be given in four sections, largely following the structure of the report.

5.1. Practical results

1. Perhaps the most obvious, but far from trivial, results of the Malmö study is that international field studies are indeed feasible. This is not to say that they are easy to arrange, for there were formidable problems of organisation to overcome when teams from ten countries conducted fieldwork simultaneously over a period of two weeks. Nevertheless, it proved to be possible. The Malmö study was almost certainly unique, but it need not be the last of its kind.

2. The design of most experiments in the field is a compromise, and this one was no exception. In theory there are many ways in which the scores of teams of observers could be compared, but in practice there are only a few. The design that was used was to have teams observing sites simultaneously; thus eight teams studied three intersections for a total period of 47 hours over nine days. This entailed having up to sixteen observers on site at a time. Some concern was expressed over the possibility of observers or teams influencing each other, but careful examination of the results showed no evidence that this occurred. Interestingly, it was the more experienced teams that expressed least concern.

3. The Malmö study has generated a unique set of data that can be used in further research at a national as well as an international level. The collection of objective data on video during the observation periods meant that, as well as comparing results between teams, each team had the

opportunity to compare their results with the objective data. This would be difficult to organise in normal circumstances, and in this way the Malmö study can make a valuable contribution to future research and development of the conflict technique.

5.2. Results of the data analysis

1. Nearly 1000 conflicts were observed by at least one team during the nine days of fieldwork; two of these were minor collisions. Cars were involved in 900 of the conflicts, cyclists in 250, pedestrians in 160, and lorries in 95.

2. Considerable differences were found among the teams in the numbers and types of conflicts they recorded. The highest scoring team recorded over four times as many conflicts as the lowest scoring team. However, it should be noted that some teams scored highly because they included numerous minor or less severe conflicts that they would not normally attach much importance to in a safety assessment exercise; thus the range in scores is not as great as may appear.

3. A multivariate analysis of the subjective scores revealed that there was a one-dimensional common structure in the data set that could justifiably be interpreted as a severity scale. On average, conflicts are scaled on this dimension in the right order by all teams. This compatibility means in effect that severity is a common concept for all teams that use the traffic conflict technique, even though their definitions and procedures might differ.

4. The outcome of this analysis was the same when carried out both on the selected set of conflicts that had been scored by four or more teams, and on the total set of conflicts scored. This implies that conflicts that are rated in total as more serious are not treated in a different way to other conflicts apart from being assigned a higher severity score. All highly scored conflicts in the total data set were present in the selected set.

5. Although this common dimension in the data is very robust, there are

still differences between the scores of the various teams. However, an important finding from the analysis is that the variations in scoring derive mainly from differences in the detection of incidents as conflicts rather than in the evaluations of severity. Once a conflict was scored, then the teams agreed to a large extent on the level of severity. Observers therefore seem to have more difficulty with the detection than with the severity rating of conflicts.

6. If the scores of the teams are related to the objective aspects of the conflict situations then it is seen that the most important variable is TTC (minimum time to collision). Conflicts with low TTC values are severe conflicts, but not all conflicts rated as severe have low TTC values. Other aspects of the conflicts that were found to be important are minimum distance between road users, conflict type and, to a lesser degree, type of manoeuvre.
In general, teams relate more to TTC than to PET (post encroachment time). The relation between TTC and conflict severity score seems to be logarithmic in form. As far as conflict type is concerned, pedestrian conflicts are regarded as the most severe and conflicts among cars and lorries as the least severe.

7. The strength of the relationship with objective data is largely influenced by the diversity in the conflict data. When dealing with homogeneous subsets of data, variables such as speed and deceleration assume greater importance. This finding has obvious implications for future research.

8. The results showed that while severity scaling is linked to objective cues, it also includes a subjective dimension due to the part played by human observers in the data collection process. Even if observers are instructed to use specific cues such as TTC or PET, they will incorporate other aspects of the situation as well. Trained observers appear to have some common understanding of conflict severity.

9. The results of the data analysis show that there is a good degree of agreement between most of the techniques calibrated in Malmö: all of them operate on the basis of a common concept. Once conflicts are detected the

level of agreement is particularly high, particularly for the more serious conflicts.

5.3. <u>Results of the team reports</u>

1. In spite of differences in the numbers of conflicts recorded, there was good agreement between teams on the proportions of conflicts classified by type of road users involved and in the proportions of conflicts recorded at the three sites studied.

2. One of the most encouraging results of the exercise was the high degree of similarity in the diagnoses made by the teams of the safety problems at the three intersections. This was of fundamental importance, since a traffic conflict technique depends as much on its ability to identify problems and interpret data as it does on the collection of appropriate data in the field. For teams from a wide range of backgrounds to be able to agree independently on the safety problems of three intersections in a small Scandinavian city is a considerale achievement.

3. The fieldwork conditions and procedures in Malmö differed from those to which several of the teams were accustomed. Nevertheless, modifications were made with little difficulty, which indicates the flexibility of existing techniques (and the teams that use them). Furthermore, the agreement among the teams demonstrates the robustness of the traffic conflict technique.

4. The prime objective of the Malmö study was to calibrate existing conflict techniques, and it was never intended that it should produce validation data comparing conflicts with accidents. However, when assessing the outcome of the study as a whole it would be somewhat perverse to ignore the accident data on the grounds that it was not directly relevant to the issue of calibration. Despite the limitations in the accident data, the results from the teams were very encouraging. Every team ranked the two sites for which reliable accident data were available in the right order, and differing by approximately the same proportions. At all three sites teams pointed to the safety problems of unprotected road users, which was very much in line with the accident data. One point

of interest was the difference between the safety problems indicated by the total number of accidents at the sites, and by those only involving personal injury. This issue could be of importance in future validation studies.

5.4. Future developments and applications

1. Conflict techniques in their present state of development are capable of being used as a complement to accident data in all fields of safety work, from research to more practical purposes. The potential applications of conflict techniques include process evaluation of safety countermeasures, behavioural analysis, safety diagnosis and countermeasure design, and more generally any investigation aimed at throwing light on the interactions between the different factors, human and technical, that contribute to danger or to safety improvement.

2. Development and application of a traffic conflict technique in countries that do not have one at present can be done in two ways: either work out a new technique, or choose from existing operational techniques the one which appears to be the most relevant to the problem in hand. In the first case, the new technique will require calibration against the others. The second solution is easier to implement as calibration results and methods for training observers are already available.

3. The calibration of conflict techniques makes it possible to promote international cooperation in traffic safety research and in the design of new safety measures: research benefits will be increased through easier transfer of results from one country to another; results should be obtained faster thanks to the use of TCT's.

4. Further international cooperation is indicated in three areas.
The first is the drawing up of guidelines for potential users. The Malmö results are sufficiently encouraging to be able to commend the traffic conflict technique to safety workers who do not at present use it. A set of guidelines for this purpose could draw on the lessons learned at Malmö, and could use the Malmö video data base for instruction purposes. The second area is that of validation. The calibration study has cleared

the way for better transfer of information and generalisation of results from individual countries. International cooperation to produce evidence capable of convincing those still sceptical about the technique is now a real possibility, and should be given a high priority.
Thirdly, it is becoming clear that the flexibility of the technique and the insight that it can provide into behavioural factors make it possible to recommend wider application of the technique into more areas of road safety research and practice. The Malmö study has shown clearly that the traffic conflict technique has a future that it can face with confidence.

REFERENCES

Amundsen, F.H.E. and Hydén, C. (1977). Proceedings First Workshop on Traffic Conflicts. Institute of Transport Economics, Oslo / Lund Institute of Technology, Oslo.

Asmussen, E. (ed) (1984). International Calibration Study of Traffic Conflict Techniques. NATO ASI Series F: Computer and System Sciences, Vol. 5. Springer-Verlag, Heidelberg.

Gifi, A. (1981). Non-linear multivariate analysis. Leyden State University, Leiden.

Horst, A.R.A. van der (1982). The analysis of traffic behaviour by video. In: Kraay (1982).

Horst, A.R.A. van der (1983). The ICTCT calibration study at Malmö: A quantitative analysis of video recordings. Institute for Perception IZF-TNO, Soesterberg.

Kraay, J.H. (1982). Proceedings of the Third International Workshop on Traffic Conflicts Techniques. R-82-27. Institute for Road Safety Research SWOV, Leidschendam.

Kraay, J.H. (1983). Review of traffic conflicts studies. Third edition. R-83-9. Institute for Road Safety Research SWOV, Leidschendam.

Leeuw, J. de and Oppe, S. (1976). The analysis of contingency tables; Log-linear Poisson-models for weighted numbers. R-76-31. Institute for Road Safety Research SWOV, Voorburg.

Older, S.J. and Shippey, J. (1980). Proceedings of the Second International Traffic Conflicts Technique Workshop, Paris, 1979. TRRL Supplementary Report 557. Transport and Road Research Laboratory TRRL, Crowthorne.

Oppe, S. (1982). Statistical tools for the calibrations of traffic conflict techniques. In: Kraay (1982).

Perkins, S.R. and Harris, J.L. (1967). Traffic conflict characteristics; Accident potential at intersections. General Motors Corporation, Warren, Michigan.

ANNEX 1. Selected conflicts, sorted on PRINCALS object scores and the classifications of these conflicts for all teams.

Confl. nr	PRINCALS score	Aus	Can	Ger	Fra	Fra	GBr	Swe	Swe	Swe	Swe	Fin	USA
502	-1.24	1	2	1	6	12	1	4	5	7	6	1	1
587	-1.24	1	2	1	6	12	1	4	5	7	6	1	1
776	-1.22	1	4	1	6	12	1	4	5	7	6	1	1
156	-1.21	1	2	1	6	12	1	4	5	7	6	4	1
539	-1.20	1	1	1	6	12	1	4	5	7	6	4	1
475	-1.18	1	4	1	6	12	1	4	5	7	6	4	1
526	-1.18	1	4	1	6	12	1	4	5	7	6	4	1
818	-1.18	1	4	1	6	12	1	4	5	7	6	4	1
781	-1.17	1	1	1	6	12	1	1	1	3	4	1	1
820	-1.16	1	4	1	6	12	1	1	1	2	3	1	1
835	-1.15	1	2	1	6	12	1	1	1	3	3	4	1
831	-1.05	1	1	1	1	04	1	4	5	7	6	4	1
916	-1.04	1	2	1	1	03	1	1	1	2	2	1	1
513	-0.99	1	4	1	6	12	2	4	5	7	6	1	1
504	-0.98	5	4	1	6	12	1	4	5	7	6	1	1
076	-0.97	1	1	1	6	12	2	4	5	7	6	4	1
658	-0.97	5	2	1	6	12	1	4	5	7	6	4	1
790	-0.97	1	1	1	6	12	2	4	5	7	6	4	1
078	-0.96	5	1	1	6	12	1	4	5	7	6	4	1
436	-0.96	1	4	1	6	12	2	4	5	7	6	4	1
440	-0.96	1	4	1	6	12	2	4	5	7	6	4	1
229	-0.93	1	1	5	6	12	1	4	5	7	6	4	1
446	-0.93	1	1	5	6	12	1	4	5	7	6	4	1
925	-0.92	1	4	1	6	12	1	4	5	7	6	2	1
811	-0.90	1	4	5	6	12	1	1	3	3	2	1	1
610	-0.87	6	4	1	6	12	1	4	5	7	6	1	1
640	-0.86	1	4	5	6	12	1	1	4	3	4	4	1
771	-0.83	1	4	1	6	12	1	4	5	7	6	1	4
778	-0.78	1	4	1	1	01	1	4	5	7	6	2	1
206	-0.75	5	4	1	6	12	2	4	5	7	6	1	1
031	-0.74	1	1	1	1	06	1	1	2	3	3	2	1
284	-0.69	5	4	1	6	12	2	1	2	3	3	1	1
059	-0.66	5	4	5	6	12	1	1	2	3	2	1	1
175	-0.64	1	1	5	6	12	6	4	5	7	6	1	1
460	-0.64	5	1	1	6	12	1	1	2	2	2	2	1
488	-0.57	5	4	1	1	04	2	4	5	7	6	4	1
144	-0.56	1	3	1	1	06	1	1	2	3	2	1	1
454	-0.55	5	2	1	6	12	1	4	5	7	6	1	2
955	-0.55	5	1	5	1	04	1	4	5	7	6	4	1
940	-0.53	1	4	1	6	12	1	4	5	7	6	2	4
621	-0.50	5	1	1	6	12	1	4	5	7	6	4	2
868	-0.50	5	4	1	1	01	6	4	5	7	6	1	1
911	-0.48	6	4	1	6	12	1	2	1	3	3	4	1

ANNEX 1. (continued)

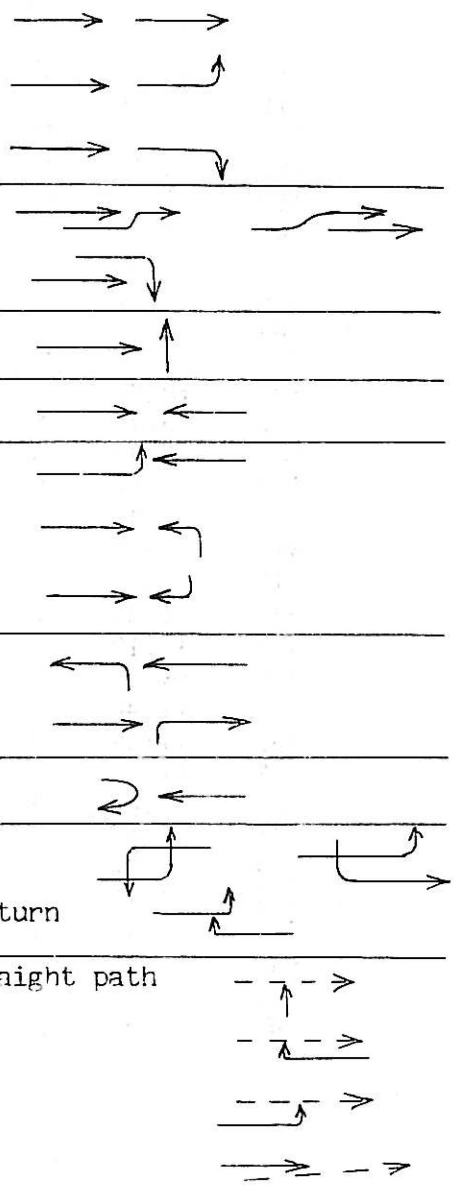
Confl. nr	PRINCALS score	Aus	Can	Ger	Fra	Fra	GBr	Swe	Swe	Swe	Swe	Fin	USA
324	-0.47	5	2	1	1	04	1	4	5	7	6	1	4
736	-0.46	1	2	1	1	02	1	2	1	4	3	2	1
243	-0.42	5	1	1	1	04	1	4	5	7	6	4	4
686	-0.37	1	4	1	6	12	6	1	3	2	2	1	2
421	-0.32	5	2	1	6	12	1	4	5	7	6	2	4
369	-0.31	1	4	1	6	12	2	4	5	7	6	2	4
432	-0.30	2	4	6	6	12	1	4	5	7	6	1	1
633	-0.29	1	2	2	1	04	2	1	2	3	3	4	1
130	-0.22	5	2	1	1	06	1	2	2	3	2	2	1
439	-0.20	1	4	1	6	12	2	2	2	2	3	1	2
279	-0.18	5	1	1	1	01	2	1	2	3	3	1	4
099	-0.17	1	3	1	1	03	1	1	3	3	3	1	4
187	-0.17	1	2	2	2	06	1	1	2	3	3	1	1
370	-0.16	1	1	2	1	04	1	1	2	3	4	1	4
615	-0.15	5	4	1	1	07	6	1	3	2	2	2	1
338	-0.14	5	2	1	1	07	6	4	5	7	6	1	4
856	-0.14	5	1	5	1	04	1	1	1	3	3	1	4
549	-0.13	1	4	5	1	04	1	4	5	7	6	2	4
089	-0.10	5	1	5	1	06	1	1	2	2	2	4	4
302	-0.08	5	2	1	2	07	1	4	5	7	6	1	4
963	-0.08	1	4	5	1	04	2	1	1	2	2	1	2
029	-0.07	1	3	1	6	12	6	1	3	3	2	2	1
906	-0.07	1	2	1	2	04	1	2	1	4	3	2	1
545	-0.05	5	4	6	1	02	1	4	5	7	6	1	2
696	-0.05	1	4	1	1	02	1	2	1	2	2	2	4
152	-0.04	5	2	1	1	01	6	1	2	3	3	4	4
477	-0.02	5	1	1	6	12	2	4	5	7	6	2	2
605	-0.02	1	2	2	6	12	2	1	3	2	3	1	2
240	-0.01	5	4	1	6	12	2	4	5	7	6	2	2
498	-0.00	2	4	1	1	07	1	4	5	7	6	4	2
239	+0.04	1	2	1	2	04	1	1	2	3	3	2	4
589	+0.05	5	3	1	1	07	1	4	5	7	6	4	4
650	+0.05	1	4	2	6	12	6	4	5	7	6	1	2
829	+0.09	5	2	5	1	07	1	4	5	7	6	2	4
376	+0.12	2	2	1	1	04	1	2	2	5	5	2	1
494	+0.15	5	3	1	1	07	2	4	5	7	6	2	1
847	+0.16	5	4	5	1	02	2	1	3	3	3	1	2
569	+0.22	6	4	5	6	12	6	4	5	7	6	4	3
624	+0.25	1	3	1	1	01	2	2	1	3	2	2	1
151	+0.28	1	4	1	1	11	6	3	3	2	3	4	4
836	+0.28	5	3	1	2	04	2	4	5	7	6	4	1
095	+0.38	1	2	2	1	01	2	2	2	3	4	4	4
030	+0.40	5	4	5	1	02	2	4	5	7	6	2	2

ANNEX 1. (continued)

Confl. nr	PRINCALS score	Aus	Can	Ger	Fra	Fra	GBr	Swe	Swe	Swe	Swe	Fin	USA
590	+0.45	5	1	5	1	07	2	1	4	3	3	2	2
218	+0.48	5	4	2	6	12	2	4	5	7	6	2	2
942	+0.55	5	4	1	6	12	2	4	5	7	6	3	2
082	+0.61	2	4	2	6	12	6	4	5	7	6	4	4
560	+0.61	2	4	2	6	12	6	4	5	7	6	4	4
266	+0.63	5	3	5	1	04	1	1	2	2	2	2	4
217	+0.64	5	1	1	6	12	6	1	3	3	2	3	4
676	+0.66	5	3	2	1	04	1	1	2	4	4	4	2
784	+0.74	6	1	5	6	12	6	3	3	5	3	4	4
321	+0.78	1	3	5	2	02	1	1	2	3	3	2	4
971	+0.82	2	4	2	6	12	2	4	5	7	6	2	2
777	+0.87	5	2	5	6	12	6	3	1	5	4	2	4
084	+0.94	5	2	2	1	06	2	2	2	4	4	2	2
023	+0.96	5	1	5	4	11	6	4	5	7	6	4	4
713	+0.96	6	1	6	2	04	2	1	1	3	3	2	4
437	+1.02	5	3	5	2	07	1	1	4	4	4	2	4
510	+1.05	1	4	2	2	02	2	2	4	4	5	2	4
035	+1.09	5	3	1	2	06	6	1	2	4	4	2	4
495	+1.12	2	3	2	6	12	1	1	2	3	2	2	2
760	+1.18	6	4	5	1	02	6	4	5	7	6	3	2
754	+1.48	5	1	5	2	02	6	3	1	3	5	2	2
967	+1.49	2	4	2	6	12	6	4	5	7	6	3	2
519	+1.63	2	3	2	6	12	2	2	4	3	3	2	3
769	+1.64	2	3	5	2	02	6	4	5	7	6	2	4
255	+1.67	2	3	6	1	04	6	2	6	3	3	2	4
349	+1.86	6	3	1	1	02	2	3	2	3	3	3	2
642	+2.16	2	3	2	2	04	2	2	2	4	5	2	2
309	+2.26	2	4	2	2	07	2	2	4	4	3	3	2
960	+3.41	5	3	2	2	02	3	2	3	5	4	3	4
675	+3.47	2	3	3	2	07	3	4	5	7	6	3	2
900	+3.83	5	3	5	3	02	3	3	1	6	4	3	4

ANNEX 2. Manoeuvre coding.

Code	Manoeuvre
1	Rear end
	Rear end with left turn
	Rear end with right turn
2	Weave or merge
	Right angle (cut-in)
3	Right angle
4	Head-on
5	Left turn
	Head-on with left turn
	Head-on with right turn
6	Right angle with left turn
	Right angle with right turn
7	U-turn
8	Double left turn
	Left turn with opposing right turn
9	Pedestrian with vehicle on straight path
	Pedestrian with right turn
	Pedestrian with left turn
	Pedestrian crossing at angle
10	Other



ANNEX 3. Participants in the Malmö study, May 30th - June 10th 1983.

AUSTRIA	Ralf Risser	Road Safety Board (Kuratorium für Verkehrssicherheit, KfV), Vienna
BELGIUM	Jozef Mortelmans	Verkeerstechniek en infrastruktuurplanning, University of Leuven
CANADA	Peter Cooper	Insurance Corporation of British Columbia, Vancouver, B.C.
DENMARK	Ulla Engel Lars Thomsen	Danish Council of Road Safety Research, Gentofte
FINLAND	Risto Kulmala Erkki Ritari Tuula Saarelma Kirsi Saulsjärvi	Technical Research Centre of Finland Road and Traffic Laboratory VTT/TIE/AUR, Espoo
FRANCE	Brigitte Baigné Dali Bouroga Nicole Muhlrad	Organisme National de Securité Routière, ONSER, Arcueil
GERMANY	Jochen Gassner Wieland Wessel Bernard Zimolong	Universität Braunschweig, Abteilung für Angewandte Psychologie
	Herbert Gstalter	TU München, Lehrstuhl für Psychologie
GREAT BRITAIN	Chris Baguley Robin Helliar-Symons Allan Wheeler	Transport and Road Research Laboratory, TRRL, Crowthorne

ISRAEL	Shalom Hakkert	Road Safety Centre, Haifa
THE NETHERLANDS	Joop H. Kraay Siem Oppe	Institute of Road Safety Research SWOV, Leidschendam
	Paul Bakker Richard van der Horst	Institute for Perception IZF-TNO, Soesterberg
SWEDEN	Sverker Almqvist Torbiörn Carlqvist Lars Ekman Christer Hydén Klas Odelid Håkan Persson Ulf Pettersson Ase Svensson	Lund Institute of Technology, Department of Traffic Planning and Engineering
	Stefan Zablocki	Malmö Gatukontor, Traffic Department
USA	William T. Baker	Federal Highway Administration, Washington, D.C.
	Jim Migletz	Midwest Research Institute, Kansas City