THE SUBJECTIVE EVALUATION OF TRAFFIC CONFLICTS BASED ON AN INTERNAL CONCEPT OF DANGEROUSNESS

HERMAN W. KRUYSE
Unit of Experimental Psychology/Centre for Safety Research, Leiden University, The Netherlands

(Received 26 June 1989)

Abstract—Earlier studies concerning subjective judgement of filmed traffic conflicts demonstrated an internal construct of dangerousness. In the present study, two experiments were carried out to investigate the construct further. Since several conflict observation techniques incorporate the impetuosity of an evasive action during a conflict, the relation between impetuosity and dangerousness was investigated. Although the impetuosity of a conflict predicts the dangerousness of a conflict rather well, it was concluded that the information underlying the judgements differed. In the second experiment the same filmed conflicts were partitioned into a commencement and a manoeuvre stage and judged on a scale of dangerousness. By relating the judgements of the different stages and the overall judgements, the information of relevance for the construct was examined. The results showed that the last part of the conflict, the part that contains most information about the minimal Time to Collision and about the Minimal Distance is of no relevance for judging the dangerousness of a conflict. It was concluded that the construct of dangerousness is based on information contained by the commencement of a convergence and by the onset of an evasive action. Finally, a conjunctive model, explaining the contribution of the two stages to the decision making process, is presented.

INTRODUCTION

A traffic location with a high incidence of accidents is called unsafe. But there is more to safety, since at this moment, we cannot predict future accident proneness of different types of traffic locations by the amount of traffic accidents in the past. Some reasons for this problem are the reliability of accident registration (Hakkert and Hauer 1988) and difficulties in defining classes of traffic locations with respect to safety (Poppe 1988; Elvik 1988).

One way to deal with these problems is the use of near-accidents, or traffic conflicts, as a measure of traffic safety. It is suggested that insight into the determinants of accidents can be achieved by analysing conflicts (Grayson and Hakkert 1987). This is based on the assumption that the process underlying conflicts is representative of the process preceding actual accidents. One could argue that the ingredients in the convergence of two or more road users resulting in either a conflict or an actual accident, correspond but are differently weighted. So, the outcome of a convergence between road users is determined by the extent to which these ingredients meet a criterion.

However, the suggestion that the registration of traffic conflicts solves the problem of predicting traffic safety is ill-founded. First, because conflicts as predictors of accidents, need to be validated; thus, the validation criterion is again accidents, and the problems involved in registration of accidents is, therefore, still not avoided. Second, because registration of conflicts based on subjective judgements, has itself a limited reliability. Thus, by using conflict instead of accident statistics, problems of inaccurate accident-reporting are exchanged for problems of subjective judgement. Consequently, the assumption that accidents are predicted by registration of conflicts is not warranted.

One way to deal with the problem of validity is to concentrate on construct validity before searching for external validity. Construct validity expresses the relation between a number of observable variables and a theoretical construct. Thus, when subjective assessment of “dangerousness” of encounters between road users is judged in a coherent manner by many subjects and across many traffic situations, it is concluded that dangerousness is a construct that underlies a substantial proportion of those judgements.
The Malmö calibration study (Grayson 1984) is an example of such an approach. The correspondence among eight different conflict techniques, used to judge traffic conflicts at three intersections in Sweden, was examined. The result of this study showed a moderate agreement between the various techniques. The best solution of a scaling analysis yielded one dimension. This dimension, labelled the “severity of conflicts,” accounted for about 40% of the variance. It was argued that the main source of variance among the techniques was their differing ability to detect conflicts, but not their scaling of severity. However an analysis of selected conflicts with a high severity score or of conflicts that were detected and judged by more than three teams, showed no better solution. Unfortunately, a plausible factor that could account for a large proportion of variance, the reliability of scoring within each team, was not investigated. After all, although the operational definitions of conflicts differ between various techniques, they all seem to have elements such as time and space and type of manoeuvre. This supports the suggestion that the techniques are substantially based on the same construct and that differences should be explained by the lack of reliability of the measurement instruments.

The development of objective tools and quantitative definitions for traffic conflict observation techniques, has in the past received much attention. The call for objective measures is, of course, inspired by the problems of subjective judgement noted above. However, objective scores will not be of much avail when the basic observations are still executed by human observers. Various so-called objective traffic conflict techniques, for example, are based on the observation of time. The minimal time to collision (MTTC), introduced by Hayward (1972), is considered as an estimator for the probability of collisions. The TTC is defined as: the hypothetical time that is required for road users to collide, if speed and path remain unchanged. An evasive action may lead to a minimum, the MTTC. The TTC was supposed to be objectively determined by sophisticated apparatus and software. Because of the enormous expense, however, human observers are used to assess conflicts based on the estimation TTC quantities. Several studies showed that the time to contact (analogous to the TTC) varying between 4 and 9 seconds, is underestimated as a function of time (Shiff and Detwiler 1979; McLeod and Ross 1983). Cavallo, Laya, and Laurent (1986) found in a field study a similar relation between TTC and time. Shiff and Detwiler observed that an objective time to contact of 2 seconds, was estimated between 4 and 9 seconds. Shinar (1984) compared subjective judgements with objective judgements of filmed traffic conflicts. The objective measure was a mathematical formulation (Balasha Hakkert, and Livneh 1980) consisting of such factors as longitudinal deceleration, radial acceleration, and approach distance. Subjects were instructed to judge the severity of conflicts in a between-groups design. In one group the severity was described as the verbal definition of Balasha's formula. In the other group severity was based on the concept of TTC. The results yielded high interrater and intrarater reliabilities but a moderate concurrent validity relative to the objective measure. It was concluded that subjects have an internal concept of near-accidents against which they consistently judge conflicts. The concept, however, does not follow the formal definitions of the severity of conflicts provided to the subjects. Kruysse and Wijlhuizen (1988) supported these findings and investigated the internal construct. It was supposed that the internal construct could be labelled as the “dangerousness” of traffic conflicts. Untrained subjects were asked to rate 39 filmed traffic conflicts on a scale of dangerousness. Some of these conflicts were taken from the Malmö calibration study. The results were represented by a one-dimensional model accounting for 59% of the total variance. Objective measures, such as the type of conflict, minimal distance, and minimal time to collision, could only partly account for the variance.

In order to generalize these findings a pilot study was conducted in the field (Kruysse 1987). All six pairwise combinations of four subjects, who participated in the above mentioned laboratory experiment (Kruysse and Wijlhuizen 1988), observed during nine days, traffic conflicts at a T-junction. The observation teams were instructed to pursue visually the cyclists and moped riders of one traffic flow from a fixed observation point and, subsequently, to judge the dangerousness of conflicts between the observed flow and all the traffic of the other flows. The interrater reliability was given by the corre-
spondence between the judgements within each pair of observers. In spite of the low number of subjects, the correspondence between subjects was in the range that could be expected on the basis of the interrater and intrarater reliabilities of these subjects in the laboratory experiment.

The results of the above-mentioned studies show that subjective judgement of dangerousness of traffic conflicts is basically reliable. It suggests that people may behave adequately in a traffic situation, because they have a clear impression of its dangerousness.

Now, where do we stand at this point in understanding the problems of reliability and construct validity? We know that subjects make use of an internal concept based on the dangerousness of conflicts. The reliability scores suggest a consistent use of strategies to assess conflicts. We also know that several sources such as the detection, the missing, and arbitrary assessment of very minor conflicts, could account for some of the unexplained variance in the field observations. It is clear, however, that the largest amount of variance, in the field and in the laboratory experiments, is expected during the assessment of the dangerousness of a conflict. Although 59% of the variance was accounted for by the between-subjects variance in the laboratory experiment, 41% remained unexplained.

The aim of the present study is to further investigate the construct of “dangerousness.” A deeper insight can be obtained when considering in more detail how people perform the dangerousness judgements. Two experiments were carried out to examine what kind of information and what kind of decision strategy subjects use in order to assess the dangerousness of conflicts.

EXPERIMENT 1

A previous experiment (Kruysse and Wijlhuizen 1988) showed that physical quantities like distance, time, type of manoeuvre, are of more or less subordinate importance in explaining the dangerousness judgements of conflicts. In evaluating the original stimuli in a sequence of increasing dangerousness, based on the results of this previous experiment, one particular aspect struck our attention. It seemed as if judgements of dangerousness were largely influenced by the impression of “control over the convergence process” by the road users involved. It was therefore hypothesized that, if the convergence appears impetuous, for example, because the action to avoid a collision by one or both road users involved is violent, the conflict is judged as dangerous. The literature shows that several traffic conflict observation techniques incorporate the aspect of impetuosity. Zimolong and Erke (1977) and Erke (1984) constructed a severity scale based on how controlled the evasive action is carried out. Güttinger (1980) also based his technique on the vigour of manoeuvres given a specific approach distance. In order to investigate the hypothesis that the dangerousness of traffic conflicts is based on the impetuosity of the manoeuvres, the next experiment was carried out. Basically, the experiment did not differ from the previous experiment. The only difference was that subjects were instructed to judge the impetuosity of the reactions between road users, instead of dangerousness.

METHOD

Apparatus

Ten monochrome video monitors were connected to one U-matic video tape recorder.

Materials

A total of 59 different video-taped traffic conflicts on intersections were selected. The values of the objective variables MTTC and Minimal Distance (MD) were calculated. The MTTC was calculated as defined by Hayward (1972) and van der Horst (1982). The MD was defined as the minimum distance between road users, as measured between
the nearest points of both road users before, during, or after the interaction. Only traffic conflicts with MTTCs between two and zero seconds were selected and proportionally divided over the stimulus set. The set contained only conflicts with opposing traffic and traffic from the side. Conflicts were possible between all types of road users including pedestrians.

**Subjects**

Thirty students of Leiden University served as paid subjects. None of them had served in a similar experiment before.

**Design**

The traffic conflicts were presented in four different sessions. The first was an introductory session in which 20 different conflicts were presented. The introductory session was followed by three different experimental sessions in each of which 13 different conflicts were presented twice, so that an experimental session contained 26 conflicts. The conflicts were presented in a random sequence, with the restriction that equal presentations were separated by at least four other conflicts. Thus, the same conflict was not repeated until at least four different conflicts had been shown. Subgroups of ten subjects were presented with the set of conflicts in a different order over the three experimental sessions.

**Procedure**

The traffic conflicts were presented on a video monitor. Subjects were instructed to judge the impetuosity of the reactions of the road users involved in traffic conflicts on a 20-point scale. The extremes were named as "not impetuous" and "very impetuous." On the score form, above the scale, a simple picture of the conflict situation was given. Thus, to avoid mistakes about the intended conflict to be judged, subjects were informed about where on the intersection and between whom the conflict occurred.

**RESULTS**

The responses were analysed using OVERALS (Verdegaal 1986, van der Burg, de Leeuw, and Verdegaal 1988), which performs a sophisticated type of factor analysis on two or more sets of variables. OVERALS transforms the variables (judgements) non-linearly into "object scores" and projects these into a multidimensional space, such that an optimal fit between the scores is generated. The optimal transformation of judgements into object scores depends on the level of measurement of the judgements. In this experiment the rating scores were treated at an ordinal level. The optimal fit is found when dimensions added to a solution result in a minimum change of stress. The maximum value of the total fit equals the number of requested dimensions if all the variance is accounted for by these dimensions. The variance accounted for by each dimension is specified by the eigenvalue.

In the first part of the analysis, OVERALS was used to determine the intrarater reliability. This was done by treating subjects as separate sets of variables and establishing a common dimension that explains a maximum proportion of the variation in the responses. Since every set in the analysis contains one variable, the OVERALS solution equals a nonlinear Principle Components Analysis. The best solution of the data was given by a one-dimensional solution. The eigenvalue of the dimension, which obviously is labelled "the impetuosity of conflicts," was 0.63. Thus, the dimension explained 63% of the total variance that reflect the intrarater reliability. The intrarater reliability, $r = 0.98$ was calculated by Pearsons Product Moment correlation between the optimal scaled responses of the first and repeated presentation.

The second part of the analysis was used to test the hypothesis that the dangerousness of conflicts is based on the impetuosity of conflicts. Thus, the correspondence between the dimension "impetuosity" and the dimension "dangerousness" obtained in a previous experiment (Kruysse and Wijlhuizen 1988), was examined. The Pearson Product Moment
An internal concept of dangerousness

The correlation between the object scores of both dimensions was \( r = 0.80 \). Since the Product Moment correlation represents a linear relation between two series of numerical data, linearity between "dangerousness" and "impetuosity," the independent variable, was tested and confirmed by relating the residuals of a regression model to the independent variable.

In order to understand more about the correspondence between both dimensions, the contribution of the objective variables MTTC and MD to the dimensions was examined using several OVERALS analyses. In this part of the analysis the objective variables were treated either as separate sets or as one set, dependent on the effect of interest. When the variables are treated as one set, the combined contribution of the variables to a dimension is examined. Further, the objective variables were categorized in different classes. The MTTC ranged between .1 seconds and 2 seconds and was divided in four classes of .5 seconds each. Likewise, the MD was divided in four classes. The first two classes of .5 meters ranged between .1 and 1 meter. The third class ranged between 1.1 meter and 2 meters. The fourth class ranged from 2.1 to 4.4 meters. These classes were based on those used in the Malmö calibration study. Table 1 shows the contribution of these variables for both dimensions.

**DISCUSSION AND CONCLUSIONS**

The one dimensional solution, which accounts for 63% of the total variance, suggests that subjects can judge the impetuosity of conflicts reliably from video. The intrarater and interrater reliabilities suggest a consistent use of assessment strategies. A regression model confirmed the assumption of linearity between the dimensions "dangerousness" and "impetuosity." "Impetuosity" seems a good predictor for "dangerousness," impetuosity accounts for 64% of the total variance. Although the correspondence between "impetuosity" and "dangerousness" is high, significant differences between the assessment strategies for both constructs appear.

The impetuosity of conflicts predicts 64% of the variance in the assessment of dangerousness. Thus, 36% remains unexplained. This means that "impetuosity" and "dangerousness" are not identical. Table 1 shows that the contribution of the combined variables to the dimension dangerousness is 33%. This is the square of the multiple correlation of the two variables within the same set and the dimension of dangerousness. The contribution of the combined variables to the dimension impetuosity is slightly higher at 37%. If the variables account for different parts of the variance, different assessment strategies should underlie the judgements of dangerousness and impetuosity. If the variables account for the same part of variance, differences in assessment strategies still exist. These differences, then, are due to the contribution of the MD, which explains two-fifth more of the variance of the dimension "impetuosity" compared to the variance that this variable explains of the dimension "dangerousness." The amount of variance explained by the MTTC is equal for both dimensions. Thus, although the differences are small between the contribution of both variables to both dimensions, it suggests at least nuanced differences in assessment strategies.

In what way then, does the dimension of impetuosity differ from the dimension of dangerousness? Is it possible that the instruction induced subjects to pay more attention to the part or phase of the conflict in which collisions are evaded? We may distinguish a number of successive stages in the development of a conflict; the beginning of the

**Table 1. The mutual and combined contributions of the experimental variables Minimal Time to Collision and Minimal Distance to the dimensions “dangerousness” and “impetuosity” (as defined by proportions of explained variance)**

<table>
<thead>
<tr>
<th>Contributors</th>
<th>Contribution to dangerousness</th>
<th>Contribution to impetuosity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MTTC</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>MD</td>
<td>16%</td>
<td>23%</td>
</tr>
<tr>
<td>MTTC + MD</td>
<td>33%</td>
<td>37%</td>
</tr>
</tbody>
</table>
convergence, the manoeuvre, and the outcome of the conflict. The dangerousness of
conflicts could be more related to the collective contribution of these sequences or phases,
while impetuosity of a conflict could be more related to the actual evasive manoeuvre.
The correlation of $r = 0.80$ indicates which weight the manoeuvering stage receives in
the judgement of the total conflict.

The contribution of successive but differently weighted stages suggest linearity be-
tween the phases. This, however, is the case only if we assume an additive relation
between the phases which, in truth, is not a necessity. So, at this stage of understanding
the assessment strategies of observers, two questions arise. First, does the information
of specific sequences or phases contribute to the overall judgement of the conflict?
Second, what kind of model describes the contribution of the phases adequately? To
answer these questions, a second experiment was conducted.

EXPERIMENT 2

In this experiment the hypothesis was tested that there are in fact different stages
that contribute separately or interactively to the judgement of dangerousness. The ex-
periment was divided into two parts that were, in fact, two experimental conditions.
Both conditions had the same filmed traffic conflicts that were used in the previous
experiments. However, in the first condition, only the beginning of the convergence (the
commencement) between road users involved in a conflict was shown. In the second
condition, the beginning of the convergence and the start of the action to evade a collision
(the manoeuvre) were shown. Thus, also in the second condition only a part of the
evasive action could be judged. The last part of the action taken and its outcome were
not presented. Hence, the commencement condition consisted of one stage, and the
manoeuvre condition consisted of two stages. Two groups of subjects, run in Condition
1 and 2, were instructed to rate the conflicts on a scale of dangerousness of the conflicts.

METHOD

Apparatus

Ten monochrome video monitors were connected to one U-matic video tape re-
corder.

Materials

In each condition, 59 different video-taped traffic conflicts at intersections were
selected. In Condition 1, the conflicts were edited in such a way, that only the Com-
 mencement of the conflict was presented. The film was cut just before the point where
the convergence between road users proceeded into an evasive action by one or both
of the road users involved. In Condition 2, the same 59 conflicts were used, but edited
differently. The film was cut just beyond the point where the manoeuvre began. This is
the point where the evasive action has started and the nature of the action can be
ascertained. Thus, the Manoeuvre includes the Commencement information and informa-
tion about the onset of the evasive action, but not of the outcome of the action. Most
conflicts under the commencement condition resulted in presentation times of about 9
seconds. No conflicts had a presentation time below 2 seconds. The commencement of
one conflict between two cyclists lasted 15 seconds. Most conflicts in the manoeuvre
condition resulted in presentation times of 1 second longer then the presentation time
in the commencement condition.

Subjects

Thirty-six students, eighteen for each part of the experiment, from Leiden Univer-
sity, served as paid subjects. None of them had participated in a similar experiment
before.

Design

The design was the same as the one used in the previous experiment.
An internal concept of dangerousness

Table 2. The Pearson Product Moment correlations between the optimal scaled responses of Condition 1 and 2 and the previous experiments

<table>
<thead>
<tr>
<th>Stage</th>
<th>1 dang*</th>
<th>1 + 2 dang</th>
<th>1 + 2 + 3 dang</th>
<th>1 + 2 + 3 imp†</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (dang)</td>
<td>-</td>
<td>0.77</td>
<td>0.78</td>
<td>0.66</td>
</tr>
<tr>
<td>1 + 2 (dang)</td>
<td>-</td>
<td>-</td>
<td>0.88</td>
<td>0.74</td>
</tr>
<tr>
<td>1 + 2 + 3 (dang)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.80</td>
</tr>
<tr>
<td>1 + 2 + 3 (imp)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*dang = dangerousness  
†imp = impetuosity

Procedure

The traffic conflicts were presented on a video monitor. Subjects were instructed to judge the traffic conflicts on a 20-point scale of dangerousness. Only the extremes were labelled— as “not dangerous” and “very dangerous.” On the score form, above the scale, a simple picture of the conflict situation was given. To avoid mistakes about the intended conflict to be judged, subjects were informed where in the intersection and between whom the conflict occurred.

Results

Considering the compatibility between the experiments, the interrater and intrarater reliabilities were analysed using OVERALS and Pearson Product Moment correlations.

For each part of the experiment (both experimental conditions) the responses were optimally represented by a one-dimensional model. The eigenvalue of the dimension in the Commencement condition was 0.54. The eigenvalue of the dimension in the Maneuuvre condition was 0.60. The correlations between the object scores of the first and repeated presentation in the first and second conditions were respectively, \( r = 0.94 \) and \( r = 0.97 \). Table 2 shows the correlations between the object scores of both conditions and previous experiments in which the two stages were not separated.

As in the Experiment 1, linearity between the variables was examined using a regression model. The assumption of linearity was not violated. A Fisher \( r \) to \( z \) transformation to these results shows that both correlations, the correlation between Stage 1 (dang) and Stage 1 + 2 + 3 (dang) \( r = 0.78 \), and the correlation between Stage 1 + 2 (dang) and Stage 1 + 2 + 3 (dang) \( r = 0.88 \), differ significantly; \( p < 0.05 \). However, the correlation between Stage 1 (dang) and Stage 1 + 2 + 3 (imp) \( r = 0.66 \), and the correlation between Stage 1 + 2 (dang) and Stage 1 + 2 + 3 (imp) \( r = 0.74 \), differed less; \( p < 0.10 \).

In order to examine the contribution of the different parts of the experiments, partial correlations between the relevant parts were calculated. These are presented in Table 3.

Discussion and Conclusions

The results demonstrate that the division of a conflict into several stages leads to some deeper insights concerning the information used for the judgement of dangerousness.

Table 3. The partial correlations between the optimal scaled responses of the experiments controlled for different conditions where (−) means partialled out

<table>
<thead>
<tr>
<th>Stages</th>
<th>( r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 + 2 (dang) − 1 (dang)</td>
<td>× 1 + 2 + 3 (dang)</td>
</tr>
<tr>
<td>1 + 2 + 3 (imp) − 1 (dang)</td>
<td>× 1 + 2 + 3 (dang)</td>
</tr>
<tr>
<td>1 + 2 + 3 (imp) − 1 + 2 (dang)</td>
<td>× 1 + 2 + 3 (dang)</td>
</tr>
<tr>
<td>1 + 2 (dang) − 1 + 2 + 3 (imp)</td>
<td>× 1 + 2 + 3 (dang)</td>
</tr>
</tbody>
</table>
The first is that the reliability of "dangerousness" judgments of conflicts is relatively independent of the amount of information contained by different stages. The interrater reliability of the judgements in Stage 1, the commencement condition, is slightly lower than the interrater reliability of the judgements in Stage 1 + 2, the manoeuvre condition. The amount of variance accounted for by one dimension in both conditions, respectively 54% and 60% deviate only to a small extent from the variance, 59% explained by one dimension in the overall judgement of dangerousness (Stage 1 + 2 + 3 in the previous experiment).

The second insight gained is that the outcome of the evasive action contributes little to nothing to the overall dangerousness judgements. The highest correlation between dangerousness judgements in different stages is produced by the linear relation between overall dangerousness judgements of the previous experiment (Stage 1 + 2 + 3 dang) and the judgements in the manoeuvre condition; \( r = 0.88 \). Hence, 78% of the variance is accounted for by Stage 1 + 2, the manoeuvre condition. This is the maximum of variance that can be obtained given the intrarater reliability of judgements in these stages. Table 3 shows that Stage 2, the correlation between Stage 1 + 2 and Stage 1 + 2 + 3 (dang) when Stage 1 is partialled out, accounts for 49% of the variance of dangerousness judgements of complete conflicts. However, the dangerousness judgements in Stage 1, the commencement condition, also correlates highly to the dangerousness judgements in Stage 1 + 2 + 3; 62% of the variance is accounted for by Stage 1. Thus almost two-thirds of the total variance of the dangerousness judgements of complete conflicts is accounted for by the commencement sequence. This is almost three quarters of the maximum variance that can be obtained given the intrarater reliabilities of the judgements in both conditions. This implies that the commencement of a conflict contains information relevant for the prediction of the dangerousness of a conflict. Hence, information about the commencement of a conflict, or information about the commencement and the nature of the action is sufficient to predict the dangerousness of conflicts.

The third insight is that impetuosity judgements are not exclusively based on the perceived evasion actions, but also on the commencement stage and the outcome of action in the manoeuvring stage. Strictly speaking, impetuosity cannot have been the basis for predicting the dangerousness of conflicts in Stage 1, since information about the evasive action and the consequences of the action were not presented in this condition. Still the correlation between dangerousness judged in Stage 1 and impetuosity judged in the complete conflict (Stage 1 + 2 + 3), was relatively high; \( r = 0.66 \). This means that the overall judgement of impetuosity is to some extent based on information other than the impetuosity of the evasive action. This is also illustrated by the drop in the correlation between the impetuosity and dangerousness of complete conflicts when Stage 1 is partialled out. Moreover, the dangerousness judgements in Stage 1 + 2 accounted for only 55% of the impetuosity judgements of complete conflicts. Whereas Stage 1 + 2 accounted for 78% of the variance of the dangerousness judgements of complete conflicts. Consequently, the overall judgements of dangerousness are better explained by the dangerousness judgements in the manoeuvre condition, than the overall judgements of impetuosity. Hence, in judging the impetuosity of complete conflicts, other aspects than the evasive action in the manoeuvre condition are involved. This may be due to the presentation of complete conflicts in Experiment 1. A total impression of the conflict was possible, although, since subjects were instructed to rate impetuosity, more attention could have been given to the evasive action in the manoeuvre sequence.

The fourth insight gained is that judgements of impetuosity are substantially affected by the last part of the evasive action and the outcome of the conflict. Table 3 shows that the correlation between dangerousness (Stage 1 + 2 + 3) and impetuosity is \( r = 0.61 \) when Stage 1 is partialled out. This correlation drops to \( r = 0.48 \) when Stage 1 + 2 is partialled out. Thus the relation between impetuosity judgements and dangerousness judgements of complete conflicts is about 23% accounted for by Stage 3; the last part of the evasive action and the outcome of the conflict. Furthermore, the correlation between Stage 1 + 2 and "dangerous" judgements of complete conflicts drops from only \( r = 0.88 \) to \( r = 0.72 \) when impetuosity judgements are partialled out.
The correlations between the judgements in different conditions with different stages are relatively high, $0.74 < r < 0.88$, with the exception of the correlation between Stage 1 and impetuosity; $r = 0.66$. The assumptions of linearity were not violated. With the exception of the commencement condition, no outliers were found. The use of natural stimuli does not always allow control over all the variables under investigation. Since the filmed conflicts were arbitrarily cut as mentioned earlier, one situation, in particular, in the commencement condition contained more information than desirable. An unsatisfactory part of the commencement sequence of that conflict was recorded. Deletion of this situation, however, did not influence the original results.

Since none of the conflicts resulted in an accident, one could argue for that reason, that the outcome of an evasive action is of less relevance for the judgement of the dangerousness of a conflict. However, the last part of the conflict and outcome of the evasive action contains information about the MTTC and PET, the Post Enchroachment Time. PET is defined as the time difference between the moment an "offending" road user leaves the area of potential collision and the moment that a "conflicted" road user arrives at the potential collision point (Allen, Shin, and Cooper 1978). The minimal PET (MPET) is the minimal time difference between these road users. Both physical quantities are used in several conflict observation techniques. Thus, the critique would apply similarly to "objective" measures based on minimal TTC and minimal PET.

The question that should be asked now is, to what kind of decision-making strategy the visual information, contained in conflicts, is submitted. The results showed that the commencement and manoeuvre sequences contain relevant information for predicting the dangerousness of complete conflicts. Table 3, however, shows that Commencement and Maneuvre are not completely correlated. Table 3 shows that when Commencement of the correlation between "manoeuvre" and the "dangerous" judgements of complete conflicts are partialled out, Maneuvre still explains 49% of the total variance of these judgements.

These results can be presented by means of a conjunctive structure (see Fig. 1): the whole conflict is judged as not dangerous only when neither the commencement stage, nor the manoeuvring stage are judged as being dangerous.

The rationale of this conjunctive model is that if one part of the conflict is dangerous, the conflict will be judged as dangerous, despite the judgement of the rest of the information contained by the conflict. If the commencement of the conflict is dangerous, the conflict is judged as dangerous. If the commencement is not dangerous but the onset of the evasive action is, in which case the action could be but is not necessarily impetuous, the conflict is also judged as dangerous. If both commencement and onset of the evasive action are not dangerous, the conflict is not judged dangerous.

Summarizing the results, the dynamic process of interaction between two road users during a traffic conflict is represented by a model with different sequences of the conflict:

<table>
<thead>
<tr>
<th>Manoeuvre</th>
<th>Commencement</th>
</tr>
</thead>
<tbody>
<tr>
<td>dangerous</td>
<td>not dangerous</td>
</tr>
<tr>
<td>danger</td>
<td>danger</td>
</tr>
</tbody>
</table>

Fig. 1. The relation between Commencement and Manoeuvre in a conjunctive structure.
as elements, and the strategies to decide whether the conflict is dangerous are represented by a conjunctive model.

GENERAL DISCUSSION

At this point it is possible to make some inferences about the perception and processing of sensory information used in order to register and to rate traffic conflicts. A remarkable finding is that the last part of the evasive action and the outcome of the action are of no relevance for the decision whether the conflict is dangerous or not. This suggests that the decision-making process is dominated by the expectancy of the outcome, rather than by the outcome itself. The study presented here showed that expectancy evolves through a comparison of information about the initial stages of a conflict with an internal concept of conflict dangerousness. The fundamental question that arises here is: how is this concept represented?

In general, expectancy is an hypothesis based on sensory input and previously learned and memorized knowledge. A red light initiates a response from a car driver that, under normal circumstances, results in a controlled braking response. The way an apprentice hits the brakes is different from that same action executed by an experienced driver. The braking action of the experienced driver will be on a relatively low level of directed control. Unless the situation asks for attention, for example the flashing lights of an oncoming car, the braking action is executed without conscious awareness (Franklin et al. 1988). It is supposed that the actual response with a low level of attention, is activated by a particular input configuration with situational context, which could be internally and externally generated (Schneider and Shiffrin 1977). Externally generated input could be physical information such as road signs, road texture, weather conditions, traffic intensity, speed and distance, and so on. Internally generated input could be the memorized road map (travelling from home to work) or the knowledge about the braking and steering characteristics of the car and so on.

Likewise, responses to complex traffic situations should be based upon internal representations of these situations. The same internal representation, can be used for judging dangerousness of filmed traffic conflicts. The assumption that the representations are linked to executive programs does not reject the use of these representations on a more abstract level.

As drivers are faced with an enormous variety of traffic situations, a large set of preprogrammed sequences of responses should be available. In order to activate these sequences a set of representations that correspond to the actual traffic situations should exist. These representations could be composed of diagrams that contain both a representation of the situation and the associated behaviour. Obviously these response diagrams are learned and continuously evaluated during the learning process. This implies, however, highly consistent responses to equal inputs. Godthelp, Milgram, and Blauw (1984) showed that time to line crossing (TLC) correlates highly with drivers accepted visual occlusion periods. TLC represents the time necessary for a vehicle to reach either edge of the driving lane. The visual occlusion periods were self-chosen and at different rates of speed. It was shown that decrease of occlusion time with higher speed corresponds to the TLC speed dependency. These findings support the assumption of internal representations.

Basically two models could account for the representation of the internal concept of conflict dangerousness. The first alternative is to represent the concept by a pattern-recognition model. Such a model suggests a rich compilation of memorized representations of complete conflicts. Visual patterns of the first stages of an actual conflict situation are processed and compared to representations of memorized conflicts. The expected outcome of the actual conflict is then based on the outcome represented by these representations. The second alternative is to represent the concept by a dynamic-equation model. Such a model consists of an algorithm in which parameters from an actual conflict situation form the input for the model. The model then performs a quickened simulation, and the outcome of this is the expected outcome of the actual
An internal concept of dangerousness

conflict. It is clear that traffic conflict observation techniques, based on the estimation of time and space quantities, originate from such a model. Compelling arguments lead to the proposition to prefer the pattern-recognition model to represent the internal construct under discussion.

The first argument is that the dynamic-equation model does not work in the commencement stage when MTTC or MPET should be estimated; the TTC and PET in the commencement stage are relatively independent of the MTTC and MPET. However, subjects are quite capable of estimating the dangerousness of conflicts at this stage. The second argument is that the dynamic-equation model applied in Stage 1 + 2 + 3 and based on MTTC and MD, poorly predicts the judgements of dangerousness of complete conflicts (Kruysse and Wijlhuizen 1988). The third argument is that people seem to process information about the conflict stochastically, whereas the dynamic-equation model suggests a continuous adjustment of parameters and a continuous availability of information about the expected outcome. The fourth argument is that the dynamic-equation model is not imitable; the model supposes a number of abilities that people do not possess. Even if people are able to select relevant parameters, then these parameters are still difficult to estimate; the estimation of accelerations is a good example. Moreover, people are not able to deal with a large set of parameters that should be estimated separately. The number of parameters mentioned before (road texture, weather and so on) amount to ten or more. In general, experimental research exhaustively showed that our visual system is highly sensitive in the recognition of visual patterns. Furthermore, the input for a pattern-recognition model does not impose a selection process for relevant parameters.

The original question in this study was to investigate the internal concept that is used to rate the dangerousness of dynamic interactions between road users. The study showed some insights into how people perform the dangerousness judgements. The most important question, however, is how these insights may contribute to the development of conflict observation techniques. The construct validity of a subjective conflict observation technique seems assured if the technique is based on the concept of dangerousness. Our study shows that people are reliable observers of traffic conflicts, if they are instructed to observe the initial stages of the conflict. Other instructions would be too complicated. Instructions based on physical quantities are only confusing and of no use. The study also showed that the impetuosity of conflicts do not sufficiently reflect the dangerousness of conflicts. This should be taken into account when observation instructions are formulated. Although this study did not clarify the nature of the 41% of variance that is not accounted for by a common dimension, construct validity could increase when hypotheses formed during observation are evaluated and corrected. Actually an analysis with two dimensions explained 67% of the total variance. The conjunctive structure of the assessment strategies of dangerousness judgements means that hypotheses formed during Stage 1 are strengthened by Stage 2, if this stage is dangerous. Thus, hypotheses are not implicitly true. However, conflict techniques, subjective or objective, are only of interest if there is a relation with traffic safety. It is clear that this relation is assumed by the existence of internal representations of dangerous conflicts.

Since these representations are supposed to represent executing programs, one should recognize that personal involvement could result in deviating judgements of dangerousness between a road user and an observer. Ergo, it may be possible that the inability to perceive personal involvement accounts for unsafe traffic behaviour. Furthermore, if subjects are more than adequate in judging the dangerousness of traffic situations, any discrepancy between judgement and behaviour should be explained by so-called human errors and technical errors. Technical errors or failures, for example blowouts, however, are underrepresented in accident data. Consequently, human errors, as a result of cognitive failure and intentional and unintentional behavior, which is prompted by respectively motivational factors like risk readiness and factors which intervene with the processing of a response, should explain traffic accidents. Risk models, like risk homeostasis (Wilde and Murdoch 1982) show some interesting explanation of unsafe traffic behavior. On the other hand, it is shown that traffic intensity contributes
H. W. Kruysse

substantially in explaining traffic accidents. Thus, factors that intervene with the processing of information and execution of an action interact somehow with traffic intensity. For example, if a road design is not what one expects from a preliminary impression, more attention could compensate for the initial behaviour. The time to compensate and the amount of attention needed to compensate, however, might be affected by traffic intensity.

The results of the experiments under discussion showed: (i) that subjects are reliable in judging subjectively the dangerousness of traffic conflicts, (ii) on what kind of information subjects base their judgement, and (iii) how subjects decide whether the conflict is dangerous or not. These results contain the constraints to the instruction of conflict observation. If these constraints are taken into account, it is possible to register subjectively traffic conflicts in order to locate, based on the frequency of the registered conflicts, dangerous spots at specific traffic locations. By testing hypotheses about the causes of these dangerous spots, insight into traffic safety could increase. On the other hand, investigation of the concept of intervening factors during the processing of relevant information suggests an interesting direction for traffic safety research.

Acknowledgements—This study was supported by the Ministry of Transport and Public Works. The author thanks Willem A. Wagenaar for his suggestions and comments, Renée Verdegaal for her assistance in using OVERALS, and A. R. A. van der Horst of the Institute for Perception (IZT—TNO) for providing the stimuli materials.

REFERENCES


Verdegaal, R. OVERALS. Department of Data Theory (UG-86-01), University of Leiden, 1986.