IZF 1990 A-40

J.E. Korteling

PROPOSAL FOR A NEW PERFORMANCE EVALUATION AND FEEDBACK SYSTEM FOR THE LEOPARD 2 DRIVING SIMULATOR

Number of pages: 33

CONTENTS

Rep.No. IZF 1990 A-40

TNO Institute for Perception,
Soesterberg, The Netherlands

Proposal for a new Performance Evaluation and Feedback system for the Leopard 2 driving simulator

J.E. Korteling

SUMMARY

In order to improve the efficiency of the training of Leopard 2 drivers the Leopard 2 driving simulator (Link Miles) is equipped with a so-called Performance And Marking (PAM) system. Application of this system may increase the objectivity of performance evaluation by the instructors and enhance the quality of behavioral feedback to the student. Preliminary usage of this system, however, indicated that the system may not be capable to realize these goals. In this connection the present report formulates seven principles concerning the design of performance evaluation and feedback systems for training simulators. These principles refer to: the validity of the simulator for different subtasks, the relevance of subtasks, the evaluation difficulty of subtasks, the relevance of performance variables, the manner of performance measurement and of criterion construction, the comprehensibility of scores, and the ergonomics of data-presentation. Based on these principles and a priory performed task-analysis (Korteling and Padmos, 1990) the PAM system is criticized and a new Performance Evaluation and Feedback (PEF) system is proposed. This new PEF system is based on a selection of the nine most relevant subtasks each combined with critical and objective performance measures. Also guidelines are provided for development, implementation, and usage of the system, calculation of scores, and data-presentation.

Rap.nr. IZF 1990 A-40 Instituut voor Zintuigfysiologie TNO, Soesterberg

Voorstel voor een nieuw Prestatie Evaluatie en Feedback systeem voor de Leopard 2 rijsimulator

J.E. Korteling

SAMENVATTING

Teneinde de efficiëntie van de opleiding van Leopard 2 bestuurders te verhogen is de Leopard 2 rijsimulator (Link Miles) uitgerust met een zogenaamd Performance and Marking (PAM) systeem. Gebruik van dit systeem kan de objectiviteit van prestatie-beoordelingen verhogen en kan de kwaliteit van actie-feedback naar de leerling toe verbeteren. Bij een eerste kennismaking met het systeem leek het echter niet aan deze toepassingen te kunnen voldoen. In verband hiermee worden in dit rapport zeven principes geformuleerd die gelden voor het ontwerpen van prestatie beoordeling en feedback systemen voor trainingssimulatoren. Deze principes hebben betrekking op: de validiteit van de simulator voor verschillende deeltaken, de relevantie van deeltaken, de evaluatie-moeilijkheid van deeltaken, de relevantie van prestatievariabelen, de prestatiemaat en de wijze van criterium constructie, de begrijpelijkheid van scores en de ergonomie van data presentatie. Op basis van deze principes en een eerder uitgevoerde taakanalyse (Korteling & Padmos, 1990) wordt het bestaande PAM systeem beoordeeld en wordt een nieuw Prestatie Evaluatie en Feedback (PEF) systeem voorgesteld. Dit PEF systeem is gebaseerd op een selectie van de negen meest relevante deeltaken met de bijbehorende kritische prestatiematen. Tevens worden richtlijnen gegeven voor ontwikkeling, implementatie en gebruik van het systeem, voor berekening van scores en voor de manier waarop data moeten worden gepresenteerd.

1 INTRODUCTION

For the training of Leopard 2 drivers a full scale driving simulator was developed by Link Miles and installed at the Dutch Driving Instruction Esquadron Tanks (RIET) in Amersfoort. In order to enhance the efficiency of instructor's task this simulator is equipped with a so-called Performance and Marking (PAM) system. Training with utilization of a PAM system provides two major advantages above normal training: additional feedback and objectivity of performance judgements[1].

Feedback is primarily relevant for the student. According to many authors (e.g. Adams, 1979, 1987; Schmidt, 1975, 1988) the training of new skills depends on the feedback ("knowledge of the results") of performed actions. Knowledge of results refers to extrinsic information about task success to the performer (Winstein & Schmidt, 1990). This enables a student to compare the outcome of his actions with the goals of actions such that he can adapt and improve his perceptual and motor skills (e.g. Adams, 1971; Schmidt, 1975). Expert instructions are intended to enhance this process indirectly by directing the student's attention to critical cues and actions (e.g. Gibson, 1967). The importance of additional action feedback has recently been demonstrated for simulator training (Lintern, Thomley-Yates, Nelson & Roscoe, 1987).

Objectivity of performance judgements is primarily relevant for the instructor who wants to get an objective representation of the strong and weak points of a student's driving behavior. Besides, objective performance data enable the instructors to improve the quality of their instructions, which in turn implies that knowledge of results (for the student) is enhanced.

On the first acquaintance with the PAM system it was noticed that, apart from technical inadequacies, the large quantity of detailed output was not easy to comprehend and seemed to lack significance for driver training. Therefore the TNO Institute for Perception was asked by the COKL of the Dutch Army to evaluate the system and to give recommendations for improvement. In the next chapter the present version of the PAM system will be described. Chapter 3 outlines problems and shortcomings of the present PAM system with respect to the training objectives. In chapter 4 a proposal will be given for development of a more appropriate and user-friendly system for performance evaluation and feedback. In chapter 5 some additional recommendations are given with reference to other devices that may improve the

---

[1]As will be seen these advantages are strongly related.

training effectiveness. In the final chapter the main conclusions and recommendations regarding improvement of the PAM system are summarized and discussed.

The Performance Evaluation and Feedback (PEF) system that will be proposed in the present report, is based on general theoretical principles (e.g. Adams, 1987, for a recent review) combined with existing knowledge of the Leopard 2 driving task (Korteling & Padmos, 1990). PEF systems for training simulators have only recently been developed and therefore knowledge concerning optimizing their efficiency is still minimally available. For example, little is known about the effects of *delay* and *frequency* of knowledge of results for complex tasks, such as vehicle driving (Swinnen, Schmidt, Nicholson & Shapiro, 1990). The present report, therefore, may be regarded as a start for improvement of the effectiveness of systems for objective performance evaluation and feedback.

2 THE PRESENT PAM SYSTEM

Feedback of the PAM system consists of a pattern of scores on predefined aspects of driving behavior related to objective criteria. In its present form the system monitors ***route driving***, consisting of terrain- and road driving, and what may be called ***obstacle driving***, i.e. water wading, driving on a low loader[2], over narrow ditches, over a "step up", over a "sloping block", or over a "knife edge". The last three obstacles all refer to a rigid object with respectively vertical sides, steep sloping sides including a traverse, and steep sloping sides without a traverse. In the present chapter both the PAM evaluation of route driving and of obstacle driving will be discussed.

Because different driving situations require different actions and/or criteria the set of performance measures, monitored by the PAM system, for route driving and obstacle driving are different. For route driving mean and/or peak values or frequencies are measured (sample rate 20 Hz). The quantities measured are listed below:

---

[2]Driving on a lowloader actually is a special action. For simplicity, however, it is subsumed under obstacle driving.

Speed, km/h (mean, maximum)
Engine rotations per minute, RPM (mean, maximum)
Acceleration/deceleration (mean, maximum)
Steering deflection as a percentage of full steering (mean, maximum)
Brake pedal deflection as a percentage of full deflection (mean, maximum)
Use of direction indicators (frequency)
Verge driving (frequency)
Smoothness of ride, based on the excursions of suspension (maximum)
Gear lever position (same vs different)
Crashes (frequency)
Slip, the tracks move across the ground at a rate which does not correspondent with the velocity of the vehicle (frequency)
Skid, the vehicle moves with blocked tracks (frequency)
Lateral deviation from criterium route (maximum)

The route driven is divided into normal (straight or curved) sections and junctions. For every single section of normal road or junction all these variables are separately measured and stored. For obstacle driving performance on the different measures is more qualitatively assessed, like very fast, good gear, hard bang to suspension, or poor heading. The following performance variables are measured:

Speed (mean, maximum)
Engine rotations per minute (mean, maximum)
Steering deflection as a percentage of full steering (mean, maximum)
Brake pedal deflection as a percentage of full deflection (mean, maximum)
Smoothness of ride, based on the excursion of suspension (maximum)
Gear lever position (same, different)
Pitch rate, the angular velocity of movement about the vehicle's horizontal axis (mean, maximum)
Heading, the direction of the vehicle's longitudinal axis relative to the longitudinal axis of the obstacle (mean, maximum)

These variables are separately measured at critical moments (e.g., first contact) of the different phases in which the obstacles are crossed. These phases are: approach, ascent, traverse, descent, and driving off (Fig. 1). Pitch rate is not measured during approach and driving off. For driving over a knife edge there is no traverse phase. Driving behavior is evaluated by relating the student's scores on a given trajectory to the results of one expert driver (the expert database) over the same trajectory. The PAM system has a facility for creating or changing these expert databases. At present 10 PAM routes of maximally 5 minutes of expert driving may be created.

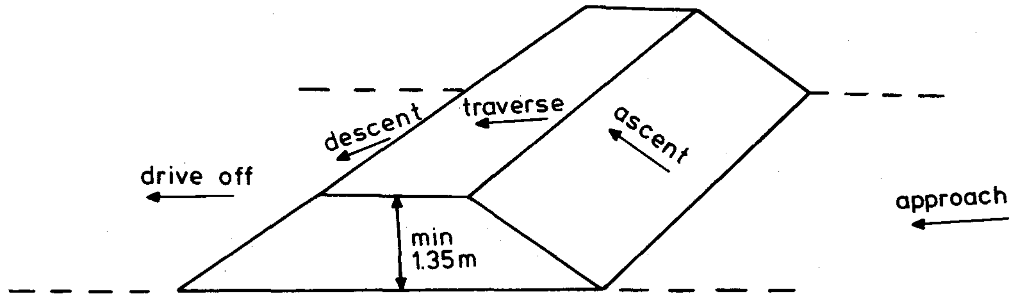

Fig. 1 Schematic representation of the five phases in which an obstacle (sloping block) is crossed.

Appendix 1 shows the heading and a partial printout of a student's driving performance (left) on a section of straight road (upper part) and across a sloping block (lower part), both related to an instructor's performance (right). With respect to route driving the student's performance on every measure is marked by the degree of closeness to the expert performance and the maximum possible mark, ranging from 2 to 12. For example, when the maximum and mean values of speed are within 5% of the corresponding expert values the student is marked 6 for each aspect. The summed score for speed will then be the maximum mark for speed, which is 12. The student is marked 3 for mean and for maximum speed when he is within 10% but beyond 5% of the corresponding expert values. When the mean and maximum speed is beyond 10% of the corresponding expert values the student will be marked 0 for speed. However, when a parameter has an average or maximum value under 10 then the marks are assigned by comparison of the expert-student discrepancy to 0.0-0.5, 0.5-1.0, and greater than 1,0. This manner of arbitrary scoring is also used for the other measures. Measures regarded as important have a higher maximum mark (e.g. speed: 12) than measures regarded as less important (e.g. gears: 4).

For each route driving measure the expert database determines the performances leading to a maximum mark. The sum of the student's marks for all measures within a section of the route (straight/curved, junction) or the obstacele (approach, ascent, traverse, descent, or drive off) is represented as a percentage of how close the student comes to the criteria represented in the expert database. Two points are added when there are no crashes detected during a section of the PAM route. The marks sum to a compound mark of 100 % when a student's driving is the same (within the minimum ranges) as the expert's driving.

The mean of all section compound percentages over a complete PAM route is called the total mark, reflecting the general similarity of the driving performance of the student relative to the instructor's driving behavior.

3 SHORTCOMINGS OF THE PAM SYSTEM

The present version of the PAM system shows problems, ranging from minor shortcomings in the clarity of the output presentation to major flaws in the selection and calculation of appropriate performance measures. The number of specific detail problems that can be identified is huge. It would take us too far to go into each particular problem. Therefore the present chapter only discusses these shortcomings on a global level. This discussion will proceed along the lines of seven principles that may be relevant for the development of systems for objective performance evaluation and additional feedback in training simulators.

1. Objective performance evaluation and additional feedback should only include subtasks that can be trained with sufficient functional validity.

The functional, or psychological, validity of a simulator (Goldstein, 1986; Baudhuin, 1987) refers to the degree to which skills learned on a simulator can be transferred to the operational system (Moraal & van Meeteren, 1990). This concept may be operationalized as the degree to which the simulator evokes the same behavior as the operational system under comparable conditions (Korteling & Padmos, 1990). In general the functional validity differs for different subtasks. Subtasks that mainly consist of procedures and/or require interaction with artificial parts of the task environment generally allow for more valid simulation than subtasks that require interaction with the natural environment (Boer, 1990; Korteling & Padmos, 1990). Van Breda and Boer (1988) showed that training the "hump area" (terrain) and the "pylon circuit" with the Leopard 2 simulator in its present form produced no transfer of skills to the operational system.[3] Increasing the objectivity and specificity of performance evaluations has no value if the skills that are evaluated differ from the skills needed on the operational system. It will thus be evident that using a PAM system for the training of these kind of subtasks only costs extra time. In general: the benefit of a PAM system for training of subtasks increases with the validity of the simulator with regard to these subtasks.

---

[3]It is expected that after the thorough rehaul planned the transfer of training will improve considerably for many manoeuvres.

Therefore the use of a PAM system should be limited to the subtasks which are simulated with sufficient validity.

2. Objective performance evaluation and additional feedback should refer to the most important and relevant subtasks of the driving task such that a broad range of skills necessary for driving performance is evaluated.

In order to maximize the efficiency of a PAM system, subtasks should be included that consume a substantial amount of training time and that cover a broad range of skills. The PAM system in its present form suggests that such a selection has not been made. For example both the knife edge and the step up are separately implemented in the PAM system, although the skills necessary for these obstacles are only slightly different. In contrast hardly any of the special actions implied in the training of Leopard 2 drivers (e.g., slalom course, vehicle clearing course) has been chosen to be monitored in the PAM system. In this way effective use of the PAM system is not optimized. In conclusion, key subtasks should be selected such that performance evaluations are valid and useful feedback is provided.

3. Objective performance evaluations should refer to subtasks which are difficult to evaluate by the instructor.

It is questionable whether the information presented by the current PAM system represents the objective information the instructor needs in order to judge driving performance. Aspects of driving behavior such as crashing or proper use of the direction indicators (which can be monitored in the instruction console) are not so difficult to evaluate that the instructor needs objective backup by a PAM system. In other words: the system should not provide information that is already evident for the instructor anyway. On the contrary, subtasks such as gas control, which are more difficult to monitor from the instruction console, should be evaluated.

4. Performance evaluation and feedback should refer to the most critical variables of subtasks.

With reference to this principle it is important to realize that the most critical aspects of a subtask may be different for different subtasks. For example, on normal straight roads driving speed will not be very important for the quality of driving performance as long as the speed remains between certain lower and upper limits. Hence, on straight roads speed control on the Leopard 2 is mainly a matter of choice (depending on other subtasks) and not one of skill. However, speed control on a Leopard 2 becomes very critical when driving in sharp curves (especially when gear "A2" is the most appropriate). This means that for different subtasks different critical variables should be chosen to represent the quality of driving performance. In order to account for this problem the PAM designers have chosen for the opposite, which may be termed a "broad spectrum approach". This means that for nearly every manoeuvre the same broad range of variables is measured. The only differentiation that has been made is the differentiation between route driving and obstacle driving. Consequently many scores that are presented give no information or give useless information concerning the subtasks involved.

5. If possible, performance measures and criteria should be defined according to objective principles, based on characteristics of the vehicle, task analysis, and formal rules for driving behavior.

In its present form performance criteria of the PAM system are based on the actual driving behavior of an instructor. The more the student scores resemble those of the instructor, the "better" his driving performance is qualified. This kind of driver evaluation is based on the assumption that for every part of a trajectory and for every variable that is measured there is one optimal value, which may be produced by any instructor. Apart from the variability of the instructors' performance, the untenableness of this assumption is demonstrated by the fact that many parts of the driving task can be performed satisfactory in many different ways. Usually there are objective limits within which the value of variables should be kept, given the driving situation, e.g. RPM 1500-2000, speed 18-30 km/h., or deceleration by braking < 3 m/s^2. There are also very simple criteria, e.g. when approaching the step up, the sloping block, or the knife edge, driving speed should be decreased until one drives at a foot-pace. Also these objects should be taken as smoothly as possible, which may be rated by measuring accelerations of the vehicle. This kind of fixed measures and criteria may easily be implemented in a new performance evaluation and feedback system, such that performance can be judged without the intermediary of an instructor.

6. Measures, scores and criteria should be easy to comprehend and implications for behavioral improvement should be clear.

With the present PAM system it is in some cases obscure what exactly is measured. For example, when the print shows mean and maximum scores on "steering" or "braking" the metrics and criteria that have been used to calculate the scores are unclear. Also the manual for the PAM system (Forber, 1986) does not specify exactly how scores and marks are measured. It was therefore not surprising that even the instruc-

tors were not always able to explain what driving behavior would lead to high scores. When it is unclear which aspects of particular actions are measured one prominent goal of the PAM system is not attained, namely: enhancing the clarity and specificity of behavioral feedback. Consequently the student still has to improve his driving performance by inefficient trial-error learning.
Secondly the prints consist of basic scores that only get meaning after comparing them to the instructor's scores. This would not be problematic if the marks, which are based on this comparison, were not weighed. This weighing only makes marks interpretable when they are related to their respective weights. The requirement of relating scores to weights in order to make them sensible makes the interpretation of scores and marks on the different PAM measures difficult and time consuming.
Furthermore, the compound marks (%) are based on a combination of (weighted) marks on all variables measured. These marks, representing the global quality of the driving behavior over a section of the total PAM trajectory, are not specific enough to draw the subject's attention to the desired cues and action sequences. They do not provide clear information concerning performance on specific subtasks or actions. The same counts *a fortiori* for the total mark, which is the mean of the compound marks taken over a complete PAM route. This total mark only may provide an objective assessment of the students global "progression" in driving skills.
In conclusion, the efficiency of the PAM system will increase substantially when it is clear to the instructor as well as to the student which aspects of driving behavior are measured. In addition, the scores and marks on prints should be directly and easily interpretable. Only when these requirements are fulfilled the system may provide useful knowledge concerning the quality of a student's driving behavior.

7. Performance data via prints should have a simple and self-explaining format; unnecessary information should not be provided.

This final principle refers to the ergonomics of data presentation. At present the data are poorly organized such that it takes a substantial amount of effort to get an overview of the performance data. A few causes of this poor ergonomics were dealt with in the priory discussed principles. For example, a clear presentation is hampered considerably by the many irrelevant behavioral data per given section (see principle 4). The problem is aggravated when a route is divided in many small parts which only take a few seconds, for each of which performance is evaluated according to the complete set of route driving variables. As a consequence of this, only one route of five minutes can produce a PAM output consisting of a few dozens of part outputs (Appendix) with a total length exceeding 1 meter.

At this point it will be evident that the PAM system in its present form violates important principles implied in the presentation of objective evaluation and performance feedback in skill training. Therefore, application of the system in its present form only takes extra time, which decreases training efficiency.

4 FUNCTIONAL SPECIFICATIONS FOR A NEW PERFORMANCE EVALUATION AND FEEDBACK (PEF) SYSTEM

The present chapter offers a selection of (aspects of) subtasks for evaluation and provides an optimal manner of measuring them according to the seven principles discussed in chapter 3.

1. Objective performance evaluation and additional feedback should only include subtasks that can be trained with sufficient functional validity.

For many subtasks the Leopard 2 driving simulator shows differences with reference to operational tracked vehicles (e.g., Padmos, 1989). Many of these differences may be corrected. According to Korteling and Padmos (1990), the differences that are suitable for correction concern the operational environment, forms and measures of the database, the dynamic vehicle-environment interaction model, and the sound system. The main structural problems are caused by the inevitable limitations of the basic concept of the simulator. This entails a computer-generated and collimated image combined with a moving base system (six degrees of freedom) governed by a computer system of limited capacity. The main consequences of this are a limited simulation of the normally available spatial and mechanical information about the natural environment and a low degree of variation and density in the simulation of other traffic. Based on a task analysis and an inventory of the structural problems of the Leopard 2 simulator Korteling and Padmos (1990) presented the following list of subtasks that probably will be trainable with sufficient effectiveness after improvement of the simulator:

Route driving

driving right on straight roads
driving left on straight roads
stopping
gearing
driving road curves
driving sharp curves and at intersections
turning on the spot

Special actions:[4]

funnel
"slalom" course
vehicle clearing course ("lane change")
parking the vehicle ("garage")
railway wagon
lowloader
driving on visual signals
driving with an image intensifier

Obstacle driving:

step up
sloping block
knife edge
small ditches (slow)
small ditches (quickly)
large ditch
camber (normal, adverse)
alternating camber
water wading

This list shows that the clusters **terrain driving** and **driving in traffic** should not be taken into consideration for implementation in a PEF system. The following principles determine what subtasks will remain on this list.

> **2. Objective performance evaluation and additional feedback should refer to the most important and relevant subtasks of the driving task such that a broad range of skills necessary for driving performance is evaluated.**

In order to use the PEF system as efficiently as possible objective evaluation and additional feedback with respect to trivial and/or overlapping subtasks has to be avoided and the range of PEF measurements has to cover the total range of driving skills as much as possible.

---

[4]The special actions were qualified as subtasks which are difficult to train effectively. One of the main reasons for this is the absence of mirrors which enable the driver to check his driving actions. However, a PAM system might provide some of the necessary feedback in order to train these subtasks more effectively.

In consultation with the instructors working with the simulator seven subtasks were qualified as trivial. Performance feedback and evaluation with reference to these subtasks was not supposed to add much to the quality of the training. Therefore these subtasks were discarded from the list above: stopping, turning on the spot, ditch, small ditches (quickly), driving on visual signals, driving with a brightness amplifier, and water wading. Primarily these subtasks require knowledge about simple procedures or actions in order to be able to perform well (see Korteling & Padmos, 1990).
There is also overlap between some of the remaining subtasks. The necessary skills for driving on a straight road (keeping a good lateral position) and gearing (choosing the right gear/speed) are largely involved in driving road curves such that both can be evaluated in a road course with curves. Furthermore, the step up, the sloping block and knife edge are comparable subtasks[5], which may be evaluated according to the same principles and procedures.
With respect to the special operations the funnel and parking the vehicle do not add much to the vehicle clearing course. In each subtask the driver has to drive between closely separated obstacles. However, only the vehicle clearing course explicitly requires the driver to make some difficult (re)positioning operations. Also between the railway wagon and the lowloader a big overlap exists. Both tasks require the driver, guided by a marshaller, to park a Leopard 2 on a transport vehicle. The lowloader is the most difficult subtask since this vehicle contains a small bump that must be taken (which also causes the marshaller to be out of sight for a moment). This implies that the railway wagon may be discarded. The following subtasks remain now on the list for a new PEF system:

Route driving
- driving right on straight sections and in curves
- driving left on straight sections
- driving in sharp curves and at intersections

Special actions:
- "slalom" course
- vehicle clearing course ("lane change")
- lowloader

Obstacle driving
- step up and steep slopes (sloping block, knife edge)
- small ditches (slow)
- camber (normal, adverse, alternating)

---

[5]In the report of Korteling and Padmos (1990) the sloping block and knife edge are referred to as "slope 60%".

3. Objective performance evaluations should refer to subtasks which are difficult to evaluate by the instructor.

All subtasks in the list above are difficult to evaluate for the instructor. The main reason for this is twofold. First, the standard instruction console belonging to the Leopard 2 simulator lacks the means to accurately monitor the position and course of the vehicle relative to its immediate environment. A "zoomable bird's-eye" monitor and a 126° driver's view presentation would enable the instructor to provide ***immediate*** feedback to the student (Korteling & Padmos, 1990). Besides, this feedback may also be given irrespective of the kind of subtask and the place in the database. Implementation of a birds-eye monitor and extra lateral monitors do not make a PEF system superfluous because of the advantages concerning objective performance evaluation, stated in chapter 1.

Table I The selected subtasks and their critical task variables.

SUBTASK	CRITICAL TASK VARIABLE
Driving right straight/curves	lateral position (steer control) correct speed/gear (gas control)
Driving left/straight	lateral position (steer control)
Sharp curves and intersections	lateral position (steer control) correct gear/speed (gas control)
"Slalom" course	lateral position (steer control) longitudinal speed (gas control)
Vehicle clearing course	lateral pos. (steer/gas control) longitudinal speed (gas control)
Lowloader	smoothness (gas/brake control) long. speed (gas/brake control) following signals of marshaller
Step up and steep slopes	smoothness (gas/brake control) long. speed (gas/brake control)
Small ditches (slow)	smoothness (gas/brake control) long. speed (gas/brake control)
Camber (adverse, alternating)	lateral position (steer control)

4. Performance evaluation and feedback should refer to the most critical variables of subtasks.

Different subtasks are based on different perceptual information and actions (task variables). The task analysis performed by Korteling and Padmos (1990) explicated the task variables implied in the different subtasks. In consultation with the instructors the most critical (important, difficult, time consuming) task variables were selected for a new PEF system. It may be expected that feedback concerning these task variables is especially useful to the student. Table I shows these critical variables for each of the selected subtasks.

5. If possible, performance measures and criteria should be defined according to objective principles, based on characteristics of the vehicle, task analysis, and formal rules for driving behavior.

When a variable is regarded useful for performance evaluation, knowledge of the vehicle and the driving task may provide unambiguous criteria and measures for specific situations. Only when sufficient knowledge about the vehicle and the driving task is lacking, such that absolute performance criteria cannot be defined, evaluation of driving behavior may be based on more relative criteria, such as expert performance. In the following only absolute performance measures and criteria will be defined for the selected subtasks.

Driving right on straight sections and in curves

With respect to lateral position the student should drive always as solid as possible on the right side of his lane (Leopard 2 vehicles are often wider than the lanes of the road) and he should not drive into the verge. The degree to which this is accomplished may be measured by separately calculating the root mean squared (RMS) error of the vehicle relative to the right edge of the road, which is not simple, and the total longitudinal distance over which the vehicle drives on the verge.

The vehicle reference point for calculating the RMS error has to be located at the longitudinal middle of the outside of the right track in the vehicle model. The criterion for RMS error is based on the idea that the right lateral armours (which laterally exceed the position of the tracks by 15 cm) should not not exeed the line marking out the right side of the road (the width of which is also 15 cm) or, in case of minor roads, should not exeed the edge of the road[6]. A high RMS error reflects a poor steering performance. By measuring the distance of verge driving instead of the duration or frequency, the speed as

---

[6] RMS error may be calculated by measuring the distance ("lane") between the tracks and the left side of the ideal line (fixed radius) marking out the right side of the road: $\sqrt{\Sigma\ (\text{lane})^2/n}$ (cm). On sections without road lines the RMS error should be calculated by measuring the distance between the right lateral armours and the edge of the road: $\sqrt{\Sigma\ (\text{lane}-15)^2/n}$ (cm). In the latter case "lane" is supposed to become negative when the track reference points drive into the verge.

well as the time of verge driving is taken into consideration[7]. The higher the speed and the longer the duration, the higher this index.
The choice of the correct speed/gear combination (gas control) only is relevant for driving in curves. Since this choice is mainly determined by the radius of the curve and the width of the road, performance may be evaluated according to the criteria presented in Table II. The duration of driving in a wrong gear has to be rated over the total PEF trajectory. This measurement should start and end when the RMS vehicle reference point (see footnote 6) enters, respectively leaves the curves.

Table II Optimal ranges for driving speed and gear positions for different curve radia (measured over the road axis) as determined by the instructors and the technical manual of the Leopard 2 (road width: 7.20 m).

RADIUS (m)	SPEED (km/h)	GEAR
10	10-15	1
15	13-18	1
20	18-23	A2
30	23-28	A2
40	25-35	A2/3
60	35-45	A3/4
80	45-60	A4
120	60-75	A4
170+	60-78	A4

Driving left on straight sections
With reference to lateral position control this subtask contains the same kind of measures as the prior one. Only, left and right have to be interchanged. Since this subtask does not involve curve driving the gear measure, representing the drivers' gas control skills, are not relevant here.

Sharp curves and intersections
The main reason for discriminating this subtask from the first one (Driving right on straight sections and in curves) is that the difficulty of maintaining the chosen gear increases substantially with the

---

[7] Vehicle reference points for boundary crossings should be placed at the outside of the right track (front and backside). The distance of verge driving over the total route, as measured at both reference points should be calculated. The mean of both values may be presented on the print.

sharpness of the curve. Since both kinds of curves constitute important successive parts of the total training course both may be considered for the new PEF system. The relevant performance measures are the same for both subtasks.

Slalom course

The slalom course consists of 10 beacons in a row, at 13 m intervals. The driver has to steer his vehicle in gear "1" around the beacons without hitting them. Since there are many ways to drive a slalom course correctly (Fig. 2) it is not possible to define an absolute criterion for lateral position that is more specific than the number of hit beacons.

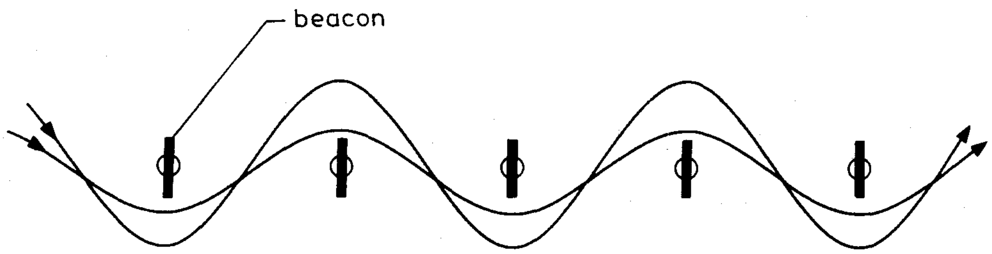

Fig. 2 Two possible manners of driving the slalom course.

As a consequence of the limited field of view in the simulator and the absence of mirrors which enable the driver to monitor his own driving behavior, intrinsic performance feedback in this subtask is very scarce. In order to enhance performance feedback to the student a clear auditive signal in the drivers cabin should indicate the moment the vehicle hits a beacon. Also the time taken to drive the course should be measured in order to represent the efficiency of driving performance. Time taken should be measured over the total slalom course starting 3 m before (frontal reference points) and ending 3 m after (backward reference points) the course.

Vehicle clearing course

The vehicle clearing course consists of one lane change to the right and one again to the original lane (Fig. 3). The vehicle clearing course should be driven in gear "A-2". The driver has to steer the vehicle as well as possible in the middle of the lanes marked out by cones. By proper gas control he also has to maintain the "A-2" gear. Task performance may thus be indicated by four absolute criteria: 1. the RMS error relative to the midline of the lanes, 2. the duration of driving in a wrong gear. 3. driving speed. In order to enhance performance feedback to the student a clear auditive signal (see slalom course) in the drivers cabin should indicate the moment the vehicle hits a cone. All measurements should start at entering the vehicle clearing course (frontal reference points) and end at leaving the course (backward reference points). RMS error can only be rated between the cones marking out the lanes. In order to take into account swerving movements of the vehicle, the reference points for RMS error calculations should be placed at the front as well as at the back in the vehicle model. The mean error of both reference points represents the quality of lateral position control.

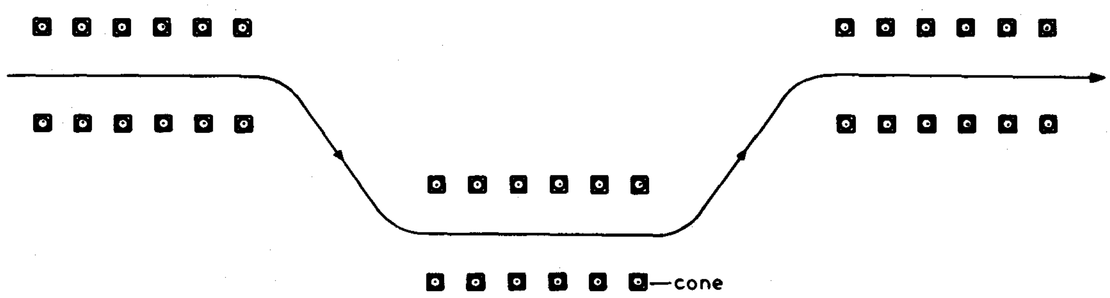

Fig. 3 Schematic representation of the vehicle clearing course.

Lowloader

A lowloader is a heavy truck designed to transport tracked vehicles (Fig. 4). Since there is just enough space for a Leopard 2 the driver has to follow signals of a marshaller when parking his vehicle on a lowloader or when driving off. Ascending as well as descending should be performed very carefully. This may be accomplished by maintaining a low driving speed and accurate brake pedal usage. When this is not appropriately done jerks may primarily be found in the acceleration profiles of the surge, heave, and pitch degrees of freedom. In order to create an optimal link between the jerks as measured and the jerks as experienced by the driver the accelerations within these dimensions have to be transformed from the center of gravity of the vehicle to the drivers place. After combining these accelerations, the derivative of this compound measure should be calculated in order to represent jerks in the driving profile. Smoothness of driving may be represented by the mean value of the three highest peak derivatives of the com-

pound acceleration[8]. The smoother a vehicle drives the lower the mean of these peak derivatives.
Also the RMS error relative to the (virtual and extended) midline of the lowloader has to be be measured. Vehicle reference points for the RMS calculations have to be located at the longitudinal middle of the vehicle model.

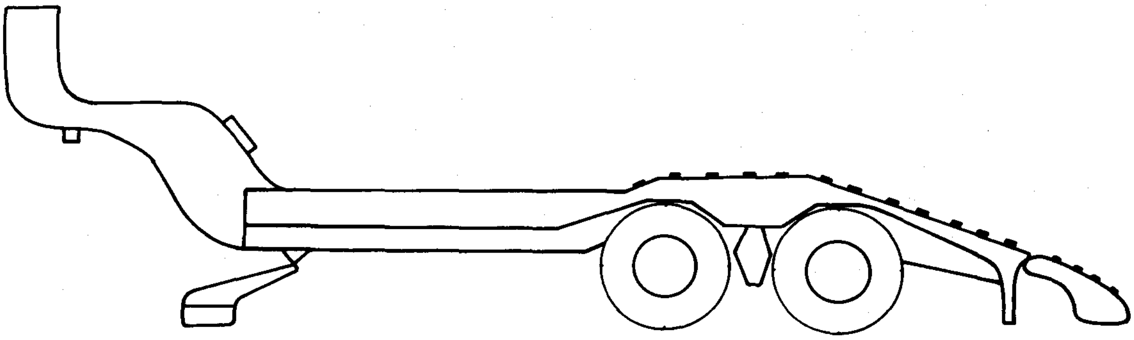

Fig. 4 Schematic representation of the lowloader.

Finally the fluency of driving behavior determines the quality of task performance. Therefore mean driving speed during this subtask should also be monitored. After having parked the vehicle on the lowloader there will be a moment of complete immobility. This period should not be taken into account for calculating mean driving speed. Speed measurements should recommence when the vehicle starts moving backwards in order to get off the lowloader. Performance registrations may start at about 5 m before the entry of the lowloader (frontal vehicle reference points), such that the straightness of approach also is taken into consideration. Measures should end when the decent is finished and drive off begins (frontal vehicle reference points).

Step up and steep slopes
Performance on crossing the step up and the steep slopes (e.g. knife edge, sloping block, Fig. 5), is mainly determined by the smoothness of vehicle movements. This means that the driver should avoided the making of jerks. Therefore for this subtask the same compound smoothness-measure may be calculated as for driving on and off the low-

---

[8] When testing the system it should be possible to change this number of peak derivatives used to rate smoothness of driving. This should also be possible for the other subtasks in which this smoothness-measure is used.

loader. The smoother a vehicle drives across an obstacle the lower the mean of the peak derivatives of the compound acceleration.
Also the fluency of crossing the obstacle determines the quality of performance. Therefore mean driving speed during this subtask should also be monitored. The measurements should begin at 2 m before the object (frontal reference points) and end at 2 m behind the object (backward reference points).

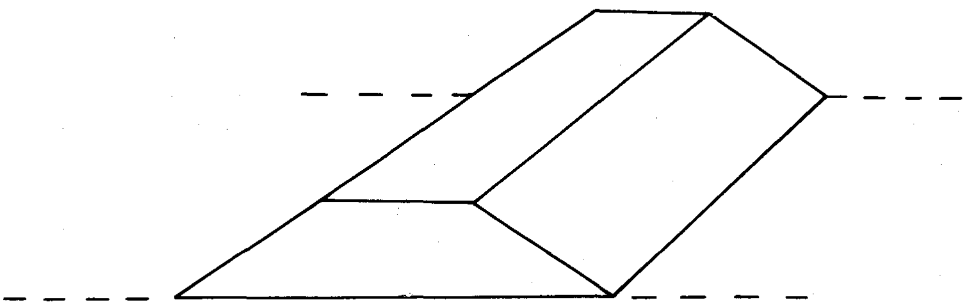

Fig. 5 Schematic representation of the sloping block.

Crossing small ditches (slow)
Also for crossing small ditches (width: 1, 2, or 3 m) the smoothness and fluency of driving are the critical performance variables. Therefore for this subtask the same compound smoothness and speed measure (using the same start and end points) may be calculated as for crossing the step up and steep slopes.

Camber (adverse, alternating)
Camber driving includes a normal camber, an adverse camber and a section with continuously changing cambers (alternating). The main problem of driving over a camber or an adverse camber is to keep the vehicle on the optimal lateral position. Therefore the RMS error relative to the right line is the best representation of task performance (see section "Driving right on straight sections and in curves"). For the alternating camber the problem is to maintain a straight and stable course by steering against continually changing lateral slopes of the road. This means that over this section just the standard deviation (relative to ones own mean lateral position) should be measured. The measurements should start and end when the frontal and backward vehicle reference points enter, respectively leave the relevant subsections.

6. Measures, scores and criteria should be easy to comprehend and implications for behavioral improvement should be clear.

Two kinds of indications of the quality of a student's performance relative to the described absolute criteria may be presented. First, simple raw scores, such as the number of cones hit or the number of gear changes. Second, transformed scores indicating the quality of driving behavior according to a certain scale (like the good old point system used at Dutch elementary schools). Raw scores provide absolute information about the concrete consequences of a student's driving actions. Transformed scores provide information concerning the level of a student's driving skills as related to the driving performances of the other students. This can be done by relating scores to the performance of other students. This relation can easily be made when scores of prior students with the same training experience are saved. This means that for every measure a database should be created in which the scores of prior students are saved. The most unambiguous transformed feedback then will be the presentation of (rounded) percentile[9] scores based on the scores of the students with the same level of prior training. The better the student's performance, the higher the percentile score, or mark.

It would be optimal to present scores on subtasks in both manners, raw as well transformed. However, with reference to RMS error and Compound smoothness the meaning of the raw scores will not be clear to most students. Therefore the latter kind of scores should only be presented after transforming them to a percentile score.

The terms for the different types of scores should be chosen such that naive people may easily understand to what they are referring. For the three task clusters: route driving, special actions and obstacles, separate total scores have to be calculated. The most obvious total score simply is the mean of the relevant percentile scores. However, the subtasks within a task cluster and the measures within a subtask are not always of equal significance. This means that the scores for the different measures have to be weighed. In consultation with the instructors weights were determined such that within each task cluster the sum of the weights was 1.0 and the individual weights reflected the relative importance of the implicated measures. Table III shows the weights for each measure within the different clusters. By adding the products of the percentile scores and their weight for all measures within a cluster the system can compute scores for the three

[9]Percentile scores can be defined as the percentage of students that had a lower score on the relevant measure. This means that the PEF system only can be used after the construction of a database of the scores of preferably at least 50 prior students.

task clusters. It was not considered necessary that these cluster scores are combined to a total score. Nevertheless it may be usefull to specify that the relative importance of route driving, special actions and obstacles is 3:2:1 respectively (see Table III). Because weighing of raw scores will influence the interpretation of total scores the weights have to be presented clearly on the printout, for example as percentages.

Table III Measures and weights for each subtask.

CLUSTER/SUBTASK		MEASURE	WEIGHT
Route driving			(3.00)
Driving left straigt	-	Optimal course	0.20
	-	Distance (m) of verge driving	0.08
Driving right straight/	-	Optimal course	0.20
curves	-	Distance (m) of verge driving	0.08
	-	Duration (s) of wrong gear	0.08
Sharp curves	-	Optimal course	0.20
and intersections	-	Distance (m) of verge driving	0.08
	-	Duration (s) of wrong gear	0.08
Special actions			(2.00)
"Slalom" course	-	Number of beacons hit	0.07
	-	Duration (s)	0.07
Vehicle clearing course	-	Optimal course	0.26
	-	Duration (s) of wrong gear	0.09
	-	Driving speed (km/h)	0.09
Lowloader	-	Optimal course	0.18
	-	Smoothness	0.18
	-	Driving speed (km/h)	0.06
Obstacle driving			(1.00)
Step up, steep slopes	-	Smoothness	0.24
	-	Driving speed (km/h)	0.08
Small ditches (slow)	-	Smoothness	0.24
	-	Driving speed (km/h)	0.08
Camber: normal/adverse	-	Optimal course	0.24
: alternating	-	Lateral stability (straight)	0.12

The cluster scores also have to be saved in a database. The databases in which scores are saved should be easy of access and it should be possible to make well documented prints of them such that their division and the discriminative power of the different measures can be studied. On a later date it should be possible to leave out measures that do not discriminate well between subjects. In that case also the weights have to be adapted, such that within a cluster their sum remains 1.00.

Table IV An example of a complete PEF print of the three task clusters as specified in the former sections. Weighed cluster scores are presented right on top of the print.

Name : W.Hoekstra	**Date:** 29.08.90	**Route Driving**	: 58
Reg. number : 580322158	**Date:** 13.09.90	**Special Actions**	: 52
Database : PEF1	**Date:** 08.10.90	**Obstacles**	: 47

TASK	**WEIGHT (%)**		**SCORE**	**MARK**
Route Driving	(50)			
Driving left straigt	20	Optimal course	--	50
	8	Distance (m) of verge driving	10	43
Driving right straight	20	Optimal course	--	69
/curves	8	Distance (m) of verge driving	32	46
	8	Duration (s) of wrong gear	63	57
Sharp curves	20	Optimal course	--	71
and intersections	8	Distance (m) of verge driving	46	55
	8	Duration (s) of wrong gear	44	45
Special Actions	(33)			
"Slalom" course	7	Number of beacons hit	4	20
	7	Duration (s)	65	51
Vehicle clearing course	26	Optimal course	--	66
	9	Duration (s) of wrong gear	1	50
	9	Driving speed (km/h)	20	40
Lowloader	18	Optimal course	--	87
	18	Smoothness	--	11
	6	Driving speed (km/h)	3	67
Obstacles	(17)			
Step up, steep slopes	24	Smoothness	--	7
	8	Driving speed (km/h)	8	62
Small ditches (slow)	24	Smoothness	--	29
	8	Driving speed (km/h)	6	51
Camber: normal/adverse	24	Optimal course	--	88
: alternating	12	Lateral stability (straight)	--	66

7. Performance data via prints should have a simple and self-explaining format; unnecessary information should not be provided.

The user friendliness and effectiveness of the PEF system will improve substantially when the output is organized such that the relevant information is provided in a clear format. The PEF output should therefore only contain information that provides clues for altering the student's driving actions in order to reach optimal (or satisfactory) driving performance. Thus a print should contain the names of the different clusters, subtasks and measures and the raw scores and marks. For each evaluation the instructor should only have to type in the date, the name of the PEF database, and the name/number of the student, which also should be found on the print. Table IV shows an example of a possible print. The data should also be accessible (by a command) immediately after one or two of the three clusters of the PEF circuit have been driven. In that case the print should not contain anything about the other clusters. When different clusters are driven on different days the print of the last day should also contain the (saved) data concerning the prior clusters. Therefore, for the different clusters, different dates may be seen on the print. The system has to be able to monitor progression of a student's driving skills. So, when a student drives the same PEF cluster several times, also the scores and marks concerning prior driving behavior on this cluster have to be presented. Of course, all text has to be presented in Dutch.

5 ADDITIONAL RECOMMENDATIONS

A PEF system as outlined above would improve the quality of performance feedback and the objectivity of performance evaluations. However, for further optimization additional measures are necessary.

First, as noted before, the instructor should have a bird's-eye monitor at his disposal, such that he always can see how the tracked vehicle is moving relative to its immediate surroundings. Only in this way the instructor can give at any moment adequate verbal information to the student about the quality and shortcomings of his driving behavior.

Second, the instruction console should also have a left and right monitor in addition to the present central monitor for watching the scene as seen through the driver's eye. By this increase in field width from 40° to 126°, instructors are enabled to evaluate steering behavior in sharp curves and at intersections. Having the disposal of a large field of view is especially important for monitoring driving during special actions and during backward driving.

The third measure concerns crash-feedback. At present the system completely breaks down at crash levels 1 to 3, which specifies serious crashes. Examples of serious crashes are: hitting a wall, a tree, a traffic participant (or another big object), or falling down from a lowloader or a step up. When serious crashes are made the image turns black and a time-consuming restarting procedure is required in order to continue the lesson. When minor crashes occur, specified by crash level 0, the vehicle drives over the object hit and continues its

course. The feedback from the crash then is more natural, which implies that mostly only a small bump is perceived. The present manner of providing crash feedback over serious and minor cases does not make sense. First, system brake down at serious crashes needlessly detains the lesson without providing effective information to both the student and the instructor. Second, natural feedback at serious crashes should be evident and thus should not need extra reinforcement. On the contrary, minor crashes are not so clear, which implies that in these cases additional artificial reinforcement could be useful. This means that in serious cases feedback should be of the same kind as natural feedback. So, when hitting the relevant objects the vehicle should suddenly stop (with a well noticeable mechanical deceleration). Also the sound system may generate a recognizable and loud natural "crash sound". Only after driving, say 3 meters backwards, it should be possible to recommence the course. This implies that the student again is confronted with the same problem. In minor cases, such as hitting a cone, a beacon, or a delineator pole normal crash feedback should be provided together with a clear auditive signal (see slalom course). This auditive signal will compensate for the lack of clear natural feedback appearing only at serious crashes. For the construction of mechanical representations of objects ("Sprites") options should be available such that an auditive signal or a sudden stop is generated when the vehicle crashes.
Finally, it may be recommended to provide extra information in the instruction console concerning strong longitudinal or vertical decelerations or accelerations.

6 CONCLUSIONS

Table V gives a summary of the main conclusions of the former sections concerning a new system for performance evaluation and additional feedback (PEF) for the Leopard 2 driving simulator. If properly implemented this system would provide a pattern of objective marks on relevant aspects of a students driving behavior, which is easily to comprehend, by the student as well as by the instructor. Besides this system would enhance the feedback to the student (knowledge of results). Apart from objective evaluation the pattern of marks would also enable knowledge of progression and of eventually persisting shortcomings in the students driving skills, such that the output may also be used for remedial teaching objectives, for example, when lessons are continued on the operational tracked vehicle.

Table V A summary of the subtasks, weights, performance measures, metrics and output presentation that should be included in a new PEF system.

Name:	**Date:** ..,..,..	**Route Driving:**	..,..,..
Registration:	**Date:** ..,..,..	**Special Actions:**	..,..,..
Database:	**Date:** ..,..,..	**Obstacles:**	..,..,..

TASK	WEIGHT (%)	MEASURE	FEEDBACK
Route Driving	(50)		
Driving left straight	20	RMS error	transf
	8	Distance (m) verge driving	raw+transf
Driving right straight/	20	RMS error	transf
curves	8	Distance (m) verge driving	raw+transf
	8	Duration (s) in wrong gear	raw+transf
Sharp curves	20	RMS error	transf
and intersections	8	Distance (m) verge driving	raw+transf
	8	Duration (s) in wrong gear	raw+transf
Special Actions	(33)		
"Slalom" course	8	Number of beacons hit	raw+transf
	8	Time needed (s)	raw+transf
Vehicle clearing course	26	RMS error	transf
	8	Duration (s) in wrong gear	raw+transf
	8	Driving speed	raw+transf
Lowloader	18	RMS error	transf
	18	Compound smoothness	transf
	6	Driving speed (km/h)	raw+transf
Obstacles	(17)		
Step up and steep slopes	24	Compound smoothness	transf
	8	Driving speed (km/h)	raw+transf
Small ditches (slow)	24	Compound smoothness	transf
	8	Driving speed (km/h)	raw+transf
Camber: normal/adverse	24	RMS error	transf
Camber: alternating	12	RMS deviation	transf

The people of the Royal Dutch Army (Workgroup Database Specifications) may create the necessary databases for the PEF system with assistance of TNO. It should be easy to link the PEF system to these databases. Furthermore it is recommended that the PEF system is made such that databases that are used for objective performance evaluation and feedback may be changed easily by the user. This means, for example, that the marshaller on the lowloader can be located anywhere. Also a well-documented manual has to be provided, containing the specifications of the system and directions for use.
Finally, the instruction console would be improved if a bird's-eye monitor and a left and right monitor were provided. Crash feedback has to be improved by natural feedback for severe crashes and additional

auditive feedback for minor crashes. When severe crashes are made the system should not brake down.

ACKNOWLEDGEMENTS

This study would not have been possible without the assistance of the people of the Royal Dutch Army working with the simulator. Therefore the author wants to thank OWI T.I.M. van de Wetering, Kpll D. Beckers, Kpll W.M. van de Bor, Kpll H. ter Haar, and Kpll L.A. Jansen for their enthusiastic help and valuable advices.

REFERENCES

Adams, J.A. (1971). A closed-loop theory of motor learning. Journal of Motor Behavior 3, 111-149.

Adams, J.A. (1979). On evaluation of training devices. Human Factors 21, 711-720.

Adams,J.A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. Psychological Bulletin 101, 41-74.

Baudhuin, E.S. (1987). The design of industrial and flight simulators. In S.M. Cormier and J.D. Hagman (Eds), transfer of learning (pp 217-237), Orlando, Fl: Academic Press.

Forber, A.S. (1986). Royal Netherlands Army Performance and Marking. Singer Link Miles.

Gibson, E.J. (1967). Principles of Perceptual Learning and Development. New York: Appleton-Century-Crofts.

Goldstein, I.L. (1986). Training in organizations: needs assessment, development, and evaluation. Monterey, CA: Brooks/Cole.

Korteling, J.E. and Padmos, P. (1990). De Leopard 2 simulator als leerhulpmiddel. Rapport IZF 1990 A-3. Instituut voor Zintuigfysiologie TNO, Soesterberg.

Lintern, G., Thomley-Yates, K.E., Nelson, B.E. & Roscoe, S.N. (1987). Content variety, and augmentation of simulated visual scenes for teaching air-to-ground attack. Human Factors 29, 45-59.

Moraal, J. & van Meeteren, A. (1990). Validatie van trainingssimulatoren. In: Militaire Spectator no. 159, 1990.

Padmos, P. (1989). Visual modelling errors in the Leopard 2 driving simulator. Report IZF 1989-6, TNO Institute for Perception, Soesterberg.

Padmos, P. (1989). Visual modelling errors in the Leopard 2 driving simulator. Report IZF 1989-6, TNO Institute for Perception, Soesterberg.

Schmidt, R.A. (1975). A schema theory of discrete motor skill learning. Psychological Review, 82, 225-260.

Schmidt, R.A. (1988). Motor control and learning: A behavioral emphasis (2nd ed). Champaign, IL: Human Kinetics.

Van Breda, L. and Boer, J.P.A (1988). Validatiestudie rijsimulator Leopard 2. Memo IZF 1988-M7, Instituut voor Zintuigfysiologie TNO, Soesterberg.

Wickelgren, W.A. (1977). Learning and Memory. Englewood Cliffs, NJ, Prentice-Hall.

Winstein, C.J., Schmidt, R.A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 677-691.

Soesterberg, December 11, 1990

Drs. J.E. Korteling

APPENDIX

NAME: H. Korteling REGIMENT: IZF-TNO REGISTRATION: 580322153 DATE: 13/01/89 INSTRUCTOR: T. van de Wetering

Database: Vlasakkers

DRIVEN OBSTACLE/ROUTE — Total mark: 63%

PROGRAMMED OBSTACLE/ROUTE — Student time: 4.39 Instructor time: 5.00

Route straight	Mark	Step: 61%			Time:0.08			Step: 100%			Time: 0.04	
	function		mean	maximum	number	mark	function		mean	maximum	number	mark
	speed	kph	14.0	26.9	x	0	speed	kph	26.1	33.1	x	12
	RPM	*100	17.72	23.37	x	6	RPM	*100	18.62	21.15	x	12
No crash 2	acceleration	%	42	75	x	6	acceleration	%	40	100	x	12
	wheel	%	x	100	2	12	wheel	%	x	100	2	12
	brake	%	14	53	x	0	brake	%	7	22	x	12
	slip	x	x	x	0	6	slip	x	x	x	0	6
	slide	x	x	x	0	6	slide	x	x	x	0	6
	deviation	mtre	x	x	x		deviation	x	x	x	x	x
	indicators	#	x	x	0.	8	indicators	x	x	x	0.	8
	verge	#	x	x	0	8	verge	x	x	x	0	8
	suspension	%		40	x	3	suspension	%		38	x	6
	gear	#	VA	V1	x	4	gear	#	VA	VA	x	4
							crash	x	x	x	x	2

Sloping Block	Obstacle mark: 55%		Time taken: 0.33		Obstacle mark: 100%		Time taken: 0.41
Approach	good speed	3.0 no brake	17 good heading	0	speed: 3.0 kph	brake: 100%	heading: 0 deg
	good gear	V2 good accel	60 good steer	0	gear: V2	accel: 59%	steering: 0
	very fast	2.4			speed at slope: 1.1 kph		
	hard bang to suspension						
Ascent	good speed	2.2 no brake	0 poor heading	358	speed: 2.3 kph	brake: 100%	heading: 351 deg
	good gear	V2 no accel	0 no steer	0	gear: V2	accel: 74%	steering: 32
	very slow	2.0 low pitch	29.59		speed at cofg: 6.5 kph		pitch rate: 41.72d/sec
	smooth ride						
Crossing	very fast	7.6 hard brake	100 aver heading	358	speed: 4.3 kph	brake: 100%	heading: 2 deg
	good gear	V2 over accel	80 good steer	0	gear: V2	accel: 76%	steering: 0
	very fast	7.5 good pitch	45.04		speed at edge: 3.6 kph		pitch rate: 46.09d/sec
	hard bang to suspension						
Descent	very fast	9.4 hard brake	100 good heading	1	speed: 5.4 kph	brake: 0%	heading: 0 deg
	good gear	V2 no accel	0 no steer	0	gear: V2	accel: 39%	steering: 32
	very fast	5.2 low pitch	38.77		speed at bottom: 4.2 kph		pitch rate: 76.28d/sec
	smooth ride						
Drive off	good speed	3.3 good brake	100 good heading	1	speed: 3.1 kph	brake: 100%	heading: 1 deg
	good gear	V2 no accel	69 good steer	0	gear: V2	accel: 80%	steering: 0
	smooth ride						

REPORT DOCUMENTATION PAGE

1. DEFENCE REPORT NUMBER (MOD-NL) TD 90-3417	2. RECIPIENT'S ACCESSION NUMBER	3. PERFORMING ORGANIZATION REPORT NUMBER IZF 1990 A-40
4. PROJECT/TASK/WORK UNIT NO. 735.3	5. CONTRACT NUMBER A89/KL/325	6. REPORT DATE December 11, 1990
7. NUMBER OF PAGES 33	8. NUMBER OF REFERENCES 16	9. TYPE OF REPORT AND DATES COVERED Final

10. TITLE AND SUBTITLE

Proposal for a new performance evaluation and feedback system for the Leopard 2 driving simulator

11. AUTHOR(S)

J.E. Korteling

12. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

TNO Institute for Perception
Kampweg 5
3769 DE SOESTERBERG

13. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Director of Army Research and Development
Van der Burchlaan 31
2597 PC DEN HAAG

14. SUPPLEMENTARY NOTES

15. ABSTRACT (MAXIMUM 200 WORDS, 1044 BYTE)

In order to improve the efficiency of the training of Leopard 2 drivers the Leopard 2 driving simulator (Link Miles) is equipped with a so-called Performance And Marking (PAM) system. Application of this system may increase the objectivity of performance evaluation by the instructors and enhance the quality of behavioral feedback to the student. Preliminary usage of this system, however, indicated that the system may not be capable to realize these goals. In this connection the present report formulates seven principles concerning the design of performance evaluation and feedback systems for training simulators. These principles refer to: the validity of the simulator for different subtasks, the relevance of subtasks, the evaluation difficulty of subtasks, the relevance of performance variables, the manner of performance measurement and of criterion construction, the comprehensibility of scores, and the ergonomics of data-presentation. Based on these principles and a priory performed task-analysis (Korteling and Padmos, 1990) the PAM system is criticized and a new Performance Evaluation and Feedback (PEF) system is proposed. This new PEF system is based on a selection of the nine most relevant subtasks each combined with critical and objective performance measures. Also guidelines are provided for development, implementation, and usage of the system, calculation of scores, and data-presentation.

16. DESCRIPTORS	IDENTIFIERS
Computer Assisted Instruction Vehicle Driving Perceptual Motor Performance Psychological Measurement Simulators Tracked Vehicles Training	Performance Feedback Performance Tests Driver Training

17a. SECURITY CLASSIFICATION (OF REPORT) -	17b. SECURITY CLASSIFICATION (OF PAGE) -	17c. SECURITY CLASSIFICATION (OF ABSTRACT) -
18. DISTRIBUTION/AVAILABILITY STATEMENT Mailing list only		17d. SECURITY CLASSIFICATION (OF TITLES) -

VERZENDLIJST

1. Hoofddirecteur van de Hoofdgroep Defensieonderzoek TNO

2. Directie Wetenschappelijk Onderzoek en Ontwikkeling Defensie

3. { Hoofd Wetenschappelijk Onderzoek KL
Plv. Hoofd Wetenschappelijk Onderzoek KL

4, 5. Hoofd Wetenschappelijk Onderzoek KLu

6. { Hoofd Wetenschappelijk Onderzoek KM
Plv. Hoofd Wetenschappelijk Onderzoek KM

7, 8, 9. Hoofd van het Wetensch. en Techn. Doc.- en Inform. Centrum voor de Krijgsmacht

LEDEN WAARNEMINGS CONTACT COMMISSIE

10. Dr. N. Guns
11. KLTZAR D. Houtman
12. Drs. C.W. Lamberts
13. Maj.Ir. G.C.A. van Langen
14. Ir. P.H. van Overbeek
15. Drs. W. Pelt
16. Maj. dierenarts H.W. Poen
17. Drs. F.H.J.I. Rameckers
18. Kol. drs. G.J.C. Roozendaal
19. LTZSD2OC KV Drs. M.B.A.M. Scheffers
20. Prof.Ir. C. van Schooneveld
21. Ir. M. Vertregt
22. Kol. vliegerarts B. Voorsluijs

23, 24. Maj. J. van der Brug, Hoofd Sie Opleidingsprojecten COKL, Afd. Plannen

25, 26. Maj. P.J.H.M. van der Burgt, Hoofd Bureau Opleidingsontwikkeling ROC

27, 28. Ing. P.J. de Haas, DMKL/OMAT

29. OWI T.I.M. van de Wetering, Beheerder Rijsimulator Leopard 2

Extra exemplaren van dit rapport kunnen worden aangevraagd door tussenkomst van de HWOs of de DWOO.
