# Objective evaluation of the quality of the FAT ES-2 dummy model

**Christian Gehre*, Sebastian Stahlschmidt#, Matthias Walz+**

* PDB – Partnership for Dummy Technology and Biomechanics, Germany

# DYNAmore GmbH, Germany

+ DAIMLER AG, Germany

**Abstract**

The numerical simulation is an essential part of the development of the passive safety of vehicles. Robust and predictable computational models are the base of the successful application of those simulations. Crash test dummies and their virtual counterparts are measuring tools to evaluate the injury risks to occupants in car crashes.

The progress of those dummy models was remarkable over the past years. By increasing the quality, the potential of further significant improvements declines. Hence, the assessment of improvements and their impact on the overall quality of simulations is getting more complex. Major improvements of sub-parts do not necessarily improve the overall performance of a model. Therefore, a standardised objective evaluation of models could ease the definition of priorities of model updates.

Objective rating tools could help to solve this problem. These tools are calculating the level of correlation between two signals, usually coming from test and simulation. All signal ratings can be merged to global ratings. However, the analysis of only one loading case is not sufficient to calculate a reliable and a robust quality score of a dummy model. A more comprehensive approach is required to provide a valid rating for all relevant loading conditions. Furthermore, it must distinguish between good and poor models and should correlate with user experiences.

This paper presents results of a study to assess the quality of the LS-DYNA FAT ES-2. The data set comprises results of dummy certification tests as well as results of various component and sled tests. The extraction of the most relevant dummy responses was an essential part of the evaluation too. Finally, all scenarios were applied to different releases of the FAT ES-2. The calculated quality scores were verified with the experiences of users of the model.

The findings of this study are limited to the FAT ES-2 model but can be transferred to another dummy model. However, the selection of loading cases and signals must be adjusted to each dummy.

**Keywords**

Objective rating, correlation of signals, dummy, ES-2, CORA

# 1 Introduction

The design of occupant safety systems by using numerical simulations became an essential part of the vehicle development processes. Especially the optimisation of safety systems as well as robustness studies of these systems benefit from the progress of the simulation. Hence, the requirements to computational dummy models increased over the past years significantly. By improving the predictability, the realisation of further improvements declines. A reliable quality rating could ease the assessment of these models.

However, increased requirements are not the only challenge. New suppliers entered the market of computational dummy models. Now there are models of different levels of detail and quality available. An objective evaluation of those models is required to find the appropriate ones.

This paper assures the findings of [6] by including results additional sled tests.

# 2 Objectives

This paper is focused on a feasibility study on the definition and application of a rating procedure to assess the level of validation of dummy models.

The check of the correct implementation of geometry, mass and inertia as well as the use of appropriate modelling techniques were not part of this study. However, they must be verified before applying any rating scheme.

# 3 Method

The application of an objective rating tool was the base of this study. It calculates the correlation of responses obtained in test and simulation. Certification tests usually cover a limited range of loads but a quality rating of a dummy model should represent almost the complete range of loading conditions. Hence, component tests as well as sled tests, conducted to validate dummy models, were included in the data set. Validation tests usually cover a wider range of loads or are more representing loads in car crashes at least.

## 3.1 Objective rating method

A standardised method with reasonable scores is the basic principle of any rating. It enables an objective and reliable assessment of the level of validation of computational models. There are a few rating tools on the market and even more published in the literature. Each of the existing tools and algorithms has pros and cons. This study used the CORA approach [5]. However, the findings of this study should be valid if another rating tool is used.

### *3.1.1 Rating tool CORA*

CORA uses two different methods to assess the correlation of signals. While the corridor method calculates the deviation between curves by using corridors, the cross correlation method analyses specific curve characteristics like phase shift or shape of the signals. The rating results ranges from 0 (no correlation) to 1 (perfect match). More information is given in [5].

### *3.1.2 Interval of evaluation*

The recording time of signals in a crash or a simulation is typically slightly longer than actually required. So the length of a signal may influence the rating. CORA offers an algorithm to extract the relevant part of the signal for the analysis. This automatism was used in all evaluations. Solely the end of the interval was set manually for some pendulum accelerations of the lumbar spine component tests. A non-relevant secondary impact of the pendulum could not be handled by the algorithm automatically [5].

### *3.1.3 Filtering of signals*

As described in [5], the chosen filter influences the rating. The analysis and assessment of smooth signals is usually more robust than the analysis of oscillating curves. So the CFC180 filter was applied to all signals.

#### 3.1.4 Preparation of the data

$T_0$ was adjusted for each test to avoid wrong rating results because of accidental phase shifts. Additionally, all data were converted to the ISO-MME format.

### 3.2 Selection of responses and weighting factors

CORA calculates the correlation of each signal separately. All single ratings were combined to a global rating by calculating the mean afterwards. Individual weighting factors are defining the significance of each signal. Those factors must be set by the user.

#### 3.2.1 Certification tests

Only the main signals were recorded in certification tests. Therefore, all evaluated signals were treated equally.

#### 3.2.2 Component, sled and vehicle tests

At first all signals were combined sensor-wise. So the sum of the weighting factors of every sensor is 1. The three abdominal forces, the three rib deflections and the three rib accelerations were treated as one sensor respectively.

Each minor axis of a sensor was assigned with a weighting factor of 0.1. A tri-axial sensor with one major axis and two minor axes is using 1x 0.8 and 2x 0.1 as weighting factors.

Finally, all sensors were combined to the total rating by using the same weighting factor for every sensor. If there were several load cases of a part or a sub-assembly available, then all load cases were treated equally.

## 4 Dummy models

The LS-Dyna FAT ES-2 model was used to demonstrate the feasibility of an objective quality rating. The quality score was calculated for three different releases of this model.

The FAT ES-2 model was developed by a consortium of German car makers and suppliers [3]. It is accepted and used all over the world.

### 4.1 Release 2.0

Version 2.0 was released in spring 2003 [4]. The model was derived from the EuroSID model, developed by the same consortium.

Additional material tests, pendulum tests with the whole dummy as well as sled tests were used to validate the model. The focus of the development was on a good overall performance of the model. In-depth validation of single parts of sub-assemblies was not in priority.

### 4.2 Release 4.5

Release 4.5 was published in summer 2009. The model was optimised by using the validation tests of release 2.0. Furthermore, the feedback of customers helped to improve the model. Compared to its predecessor, the numerical robustness of this release increased significantly.

### 4.3 Release 5.0

It was decided by PDB in 2009 to start a major update of the ES-2 model to improve the quality. Therefore, new material tests, component tests and sled tests were defined and conducted. The focus was on the improvement of the most crucial parts of the existing model such as shoulder, abdomen and lumbar spine. Particular attention was paid on the component and the sub-assembly level. Version 5.0 of the LS-Dyna model was released in spring 2011. The model used in this study is not the finalised version 5.0 but very close to the final release.

In principle, this release can be compared to version 2.0 of the model. It is the first version after the completion of a new development or a major update. The full potential of the new test data will probably be realised with the successor of version 5.0

## 5 Load cases

The quality of the three different releases was assessed by using certification, component and sled tests. Simulation runs within a vehicle environment were used as a final proof of the findings.

### 5.1 Certification tests

All certification tests of the ES-2 are described in [2]. Different test set-ups check the conformity of head, neck, shoulder, ribs, abdomen, lumbar spine and pelvis with the specs.

The assessment of head and neck was not included in this study. The focus was on thorax and pelvis.

#### *5.1.1 Shoulder*

The longitudinal acceleration of the pendulum was used as the only signal to calculate the quality score.

41 dummy certifications were the basis of the evaluation.

#### *5.1.2 Thorax*

The performance of the three ribs is tested in single rib tests. A pendulum impact against the complete thorax but without arm can be applied alternatively [1] which was used in this study. The deflection of the ribs as well as the longitudinal pendulum acceleration was assessed.

The data set used is less extensive compared to the other certification tests. Only four tests of two dummies were available.

#### *5.1.3 Abdomen*

The abdomen is certified in a pendulum test. Usually, the sum of the three abdominal load cells is evaluated. To get a more reliable rating of the abdomen, the three abdominal forces were assessed separately. The pendulum acceleration completes the set of evaluated signals.

In total 41 certifications of three different dummies were used as base for all evaluations.

#### *5.1.4 Lumbar spine*

The lumbar spine is tested in a pendulum test with a mass substitute mounted on top. Three different bending angles were assessed.

40 certification tests coming from four specimens were taken as basis for the evaluation of the model.

#### *5.1.5 Pelvis*

A pendulum test is used to certify the pelvis. The pubic force as well as the longitudinal pendulum acceleration was taken for the assessment of the correlation.

The reference data set included 40 certification tests of three different dummies.

### 5.2 Component tests

The programme to update release 4.5 of the ES-2 model started with extensive dynamic tests of several dummy parts and sub-assemblies. This study used pendulum tests with clavicle, abdomen and lumbar spine for the evaluation.

#### *5.2.1 Clavicle*

The clavicle was fixed via shoulder load cell to the test rig and was loaded by a pendulum in different directions and with different energies. Figure 1 shows the test set-up of the vertical impact to the clavicle exemplarily.

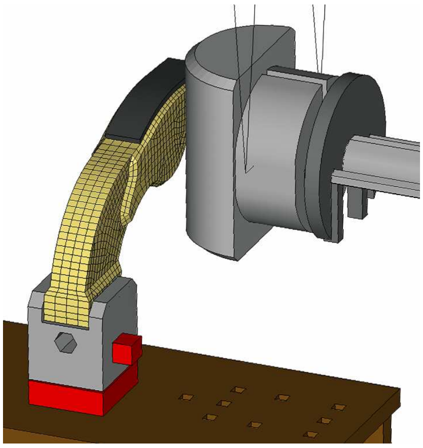

Figure 1. Impact z to the clavicle.

The forces of the shoulder load cell as well as the longitudinal and transverse accelerations of the pendulum were taken for the assessment of the correlation between test and simulation.

#### *5.2.2 Abdomen*

Figure 2 shows the set-up of one abdomen pendulum test. The abdomen was loaded with different energies, at different impact locations and impact angles. The focus of the validation work was on the pure lateral impacts. So the weighting factors of the oblique tests were reduced.

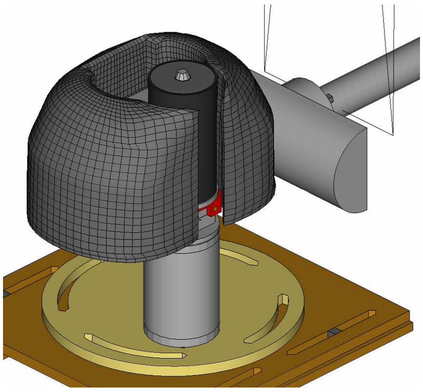

Figure 2. Lateral impact to the abdomen.

The three abdominal forces and pendulum accelerations were taken for the assessment.

#### *5.2.3 Lumbar spine*

Three set-ups were used to identify the properties of the lumbar spine for pure torsion, shear and flexion bending loads. All modes were tested with different impact energies.

The signals of the pendulum, the T12 and the lumbar spine sensors were taken for the analysis of the shear and the bending mode. The moment about the vertical axis of the lumbar spine was the only signal of the assessment of the torsion tests.

### 5.3 Sled tests

Results of two test series were taken to assess the quality of the models. The PDB tests were conducted to update the ES-2 release 4.5. The more extensive FAT tests were used to develop the first release of the ES-2 model.

Sled tests with rigid bench and rigid barriers were used to validate the global kinematics of the dummy as well as the interactions between sub-assemblies. The different barrier faces induce kinematics and loadings observed in various vehicle crashes. Each barrier was assigned with a specific code (D1, D2 etc.) to differentiate between them. The indices “P” and “F” are used to differentiate between PDB and FAT test series. The geometry of the barriers used in both series is almost identically. However, some details may differ.

#### *5.3.1 PDB tests*

Tests with the $D1_P$, $D3_P$ and $D4_P$ barriers were used in this study. $D1_P$ and $D4_P$ are flat barriers. The upper edge of $D4_P$ is at the same level like the upper rib of the dummy, whereas the $D1_P$ barrier covers the whole shoulder. The $D3_P$ barrier is very similar to $D1_P$ but is equipped with an additional rigid pelvis pusher.

The dummy was placed on the WorldSID bench in all tests. Additional information is given in [9]. Figure 3 shows the set-up of a test with the $D3_P$ barrier exemplarily.

Head and neck loads were again not considered in the evaluation but the following signals were taken to calculate the correlation with the test data:

- Acceleration of T1, T12, ribs and pelvis
- Forces and moments of shoulder, T12, abdomen and pelvis
- Deflection of the ribs

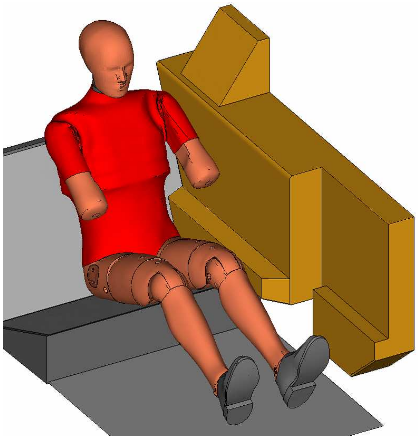

Figure 3. Sled test with $D3_P$ barrier.

#### *5.3.2 FAT tests*

The FAT set-up differs from the PDB tests. The geometry of the rigid bench was derived from the ECE-R16 seat.

The number of sensors was less comprehensive compared to the PDB sled test programme. So the assessment is based on major loads only. The following signals were taken for the quality rating:

- Acceleration (y) of T1, ribs and pelvis
- Forces (y) and moments (x) of T12, abdomen and pelvis
- Forces (x, y) of the shoulder
- Deflection of the ribs

In total nine configurations were used for the validation of the dummy and finally, for the assessment of the dummy. The sample comprises six barriers and two impact velocities.

### 5.4 Vehicle tests

A good validation in certification, component and sled tests is the base for any successful application of the model in vehicle development processes. However, it is not a guarantee for high predictability in vehicle simulations. Additional simulation runs with a vehicle environment consolidate the findings.

A 90° pole impact of a mid-size vehicle without deployed side airbag and fired belt pre-tensioner was used as reference test. The selection of signals and the corresponding weighting factors was taken from the evaluation of the PDB sled tests.

## 6 Results of the rating

The absolute classification of the CORA rating is complicated. Ratings close to 1 are easy to understand – the correlation is almost perfect. Unfortunately, many ratings are between 0.5 and 0.7. It is not clear yet, when a rating represents a good model. The significance of differences between ratings rises with the absolute deviation. However, a relative assessment by using the rating of one model as reference is preferred at this time.

The CORA algorithm is already used for a couple of years. Based on experiences made, some assumptions can be given. A good correlation can be assumed if the rating of a single signal is clearly better than 0.8. The situation is more complicated in case of assessing a complete test of numerous signals. Correlations with a score of 0.7 or higher could be assumed as good.

### 6.1 Certification tests

Certification tests are part of most of the dummy validation programmes. The focus usually is to meet the requirements (e.g. corridors) of every test and not on an overall good correlation of the responses. This information might explain the rating results of some certification tests.

The results of the CORA rating of the dummy certifaction tests is shown in Table 1. Almost all tested body segments from release 2.0 to 5.0 were improved significantly. The new test data used for the ES-2 update programme enabled a more profound validation. Thorax and shoulder of version 5.0 correlate very well to the hardware in this specific set-ups.

The limited improvement and even partly loss of correlation of release 4.5 compared to 2.0 is probably based on the development process of this version. As mentioned above, both releases used the same validation data set. Version 4.5 was mainly optimised for load cases in vehicle environments. So a loss of quality in some certification tests was an accepted side effect.

Table 1.
Evaluation of certification tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| Shoulder | 0.562 | 0.645 | 0.825 |
| Thorax | 0.841 | 0.919 | 0.911 |
| Abdomen | 0.532 | 0.576 | 0.774 |
| Lumbar spine | 0.394 | 0.397 | 0.568 |
| Pelvis | 0.748 | 0.625 | 0.785 |

### 6.2 Component tests

Table 2 shows the assessment of clavicle, abdomen and lumbar spine. Only release 5.0 was validated against those tests. Consequently, its score is better than that of the previous model releases.

Table 2.
Evaluation of component tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| Clavicle | 0.551 | 0.594 | 0.776 |
| Abdomen | 0.690 | 0.714 | 0.750 |
| Lumbar spine | 0.675 | 0.562 | 0.731 |

Anyhow, the rating indicates that the lumbar spine seems to be a crucial dummy part. By increasing the overall dummy performance of release 4.5, the quality of the lumbar spine decreased significantly. The lumbar spine of version 5.0 shows good correlation but the good results of the component tests seem to be in conflict with the rating of the certification tests.

### 6.2.1 Clavicle

Detailed information of the evaluation of the clavicle is given in Table 3. The total rating of each impact (impact x, y and z) is calculated from the evaluation of two sub-load cases.

Table 3.
Evaluation of clavicle tests

| | R2.0 | R4.5 | R5.0 |
|---|---|---|---|
| Impact x | 0.636 | 0.635 | 0.619 |
| Impact y | 0.752 | 0.820 | 0.793 |
| Impact z | 0.681 | 0.687 | 0.837 |
| Mean | 0.551 | 0.594 | 0.776 |

As mentioned before, the improvements of release 4.5 are achieved by optimising the clavicle without new component tests. So the progress is limited. Surprisingly, release 5.0 does not benefit from the new test data in longitudinal and lateral loadings. Solely the correlation of vertical impacts increased significantly.

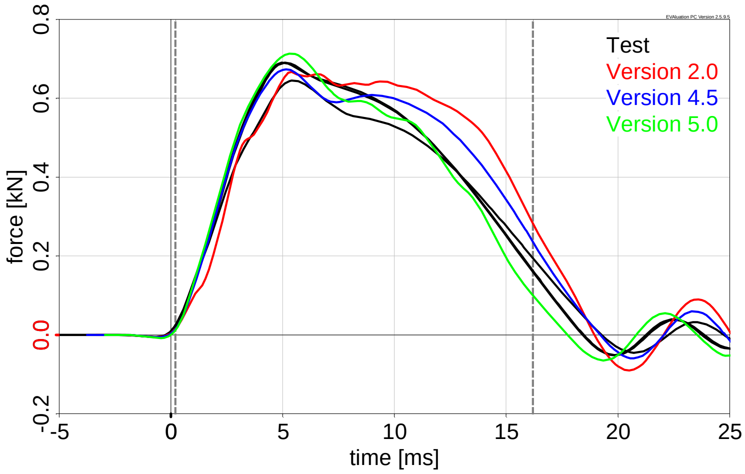


Figure 4. Impact y – shoulder force $F_y$.

Figure 4 shows the lateral shoulder forces of an impact y test exemplarily. The corresponding CORA rating is shown in Table 4. The vertical dashed lines visualise the evaluated interval of the signals. In spite of the big differences of the CORA rating, the signals of the three models are close to the test data. Therefore, small differences (approx. <0.05) of the CORA rating should not be overestimated.

Table 4.
Evaluation of the shoulder force

| | R2.0 | R4.5 | R5.0 |
|---|---|---|---|
| Shoulder force (impact y) | 0.710 | 0.878 | 0.963 |

### 6.2.2 Abdomen

A summary of the abdomen tests is shown in Table 5. The weighting factor of the perpendicular impacts is 0.333 and 0.166 of the oblique tests. The quality score of both 90° configurations is calculated from six sub-load cases respectively. Each oblique impact represents only one load case.

Table 5.
Evaluation of abdomen tests

| | R2.0 | R4.5 | R5.0 |
|---|---|---|---|
| 90°, mid pos. | 0.626 | 0.549 | 0.843 |
| 90°, upper pos. | 0.400 | 0.567 | 0.784 |
| 60°, mid pos. | 0.639 | 0.612 | 0.782 |
| 120°, mid pos. | 0.614 | 0.719 | 0.618 |
| Weighted mean | 0.690 | 0.714 | 0.750 |

The new tests helped to improve the quality of the abdomen of release 5.0 under pure lateral load remarkably. The current state of version 5.0 is validated by using the pure lateral impacts only. So the oblique impacts could be used for further improvements.

The good correlation of release 4.5 in impacts at 120° is probably a side effect of the optimisation. In-depth analysis shows that almost all signals of this model correlate slightly better with the tests. Finally, the numerous minor improvements result in a good overall rating.

#### *6.2.3 Lumbar spine*

Table 6 shows more detailed information on the assessment of the lumbar spine. Each rating is a combination of the assessment of three sub-load cases.

Table 6.
Evaluation of lumbar spine tests

| | R2.0 | R4.5 | R5.0 |
|---|---|---|---|
| Flexion | 0.735 | 0.709 | 0.844 |
| Shear | 0.685 | 0.617 | 0.899 |
| Torsion | 0.606 | 0.306 | 0.450 |
| Mean | 0.675 | 0.562 | 0.731 |

Flexion and shear of version 5.0 improved significantly by using the new test data. However, pure torsion seems to be a problem of version 4.5 and 5.0.

### 6.3 Sled tests

Table 7 shows the non-weighted mean rating of the two sled test series. The rating of the both tests series is very similar. Solely version 2.0 of the dummy lost some scores due to a not completed simulation run.

Table 7.
Evaluation of PDB and FAT sled tests

| | R2.0 | R4.5 | R5.0 |
|---|---|---|---|
| PDB tests | 0.390 | 0.579 | 0.666 |
| FAT tests | 0.552 | 0.592 | 0.668 |

#### *6.3.1 PDB tests*

Table 8 gives an overview on the results of the evaluation of the PDB barrier tests. The ES-2 version 2.0 could not be assessed with barrier $D4_P$ because of numerical instabilities of the rib damping material. Consequently, the score was set to 0.

Table 8.
Evaluation of PDB sled tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| $D1_P$ barrier | 0.536 | 0.509 | 0.617 |
| $D3_P$ barrier | 0.634 | 0.616 | 0.724 |
| $D4_P$ barrier | 0.000 | 0.612 | 0.657 |

The dummy responses of model release 5.0 correlate clearly better with the test data than those of the previous model releases. In-depth analysis showed that the quality of almost all dummy parts is improved.

The rating of version 4.5 is remarkable. The improvement of the model's robustness and the tuning of the performance of sub-assemblies reduced the correlation of the complete model in those sled tests. However, it should be considered that the assessed tests were not part of the validation programme of release 2.0 and 4.5.

### *6.3.2 FAT tests*

The FAT sled tests were an essential part of the continuous model improvement up to model release 4.5. Therefore, the progress of version 4.5 compared to version 2.0 is more remarkable than in the PDB tests. However, the evaluation of the FAT sled tests confirms the findings of the PDB sled tests mostly. All ratings of the FAT sled tests are shown in Table 9.

Table 9.
Evaluation of FAT sled tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| $D1_F$ barrier, v1 | 0.587 | 0.652 | 0.775 |
| $D1_F$ barrier, v2 | 0.537 | 0.590 | 0.587 |
| $D3_F$ barrier, v1 | 0.574 | 0.560 | 0.695 |
| $D3_F$ barrier, v2 | 0.634 | 0.600 | 0.697 |
| $D4_F$ barrier | 0.479 | 0.602 | 0.667 |
| $D5_F$ barrier | 0.545 | 0.561 | 0.657 |
| $D6_F$ barrier, v1 | 0.516 | 0.525 | 0.640 |
| $D6_F$ barrier, v2 | 0.634 | 0.723 | 0.794 |
| $D7_F$ barrier | 0.464 | 0.514 | 0.497 |

## 6.4 Vehicle tests

The results of the quality rating of a vehicle test is shown in Table 10. The ranking of the models is identical to that of the sled tests. Solely the absolute difference between release 4.5 and 5.0 is reduced.

Table 10.
Evaluation of a vehicle test

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| Vehicle | 0.655 | 0.671 | 0.739 |

### 6.5 Influence of signal weighting factors

The definition of the weighting factors of the signals has got an influence on the total rating. Table 11 shows the rating of the lumbar spine by using the same weighting factor for all major and minor signals. These results should be compared to the regular rating shown in Table 6.

Table 11.
Alternative rating of lumbar spine tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| Flexion | 0.769 | 0.758 | 0.837 |
| Shear | 0.776 | 0.722 | 0.898 |

Uniform weighting factors improve the rating of the lumbar spine. The ranking between the models remains the same. It is an indication that reasonable weighting factors generate reasonable ratings. So the rankings shown in this study are valid.

Table 12.
Alternative rating of PDB sled tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| $D1_P$ barrier | 0.473 | 0.467 | 0.533 |
| $D3_P$ barrier | 0.517 | 0.534 | 0.620 |

Uniform weighting factors worsen the rating of the PDB sled tests (see Table 12 and Table 8). Many minor signals of poor correlation got more influence on the total results. However, the general tendencies of the regular rating are confirmed. Solely release 2.0 and 4.5 are switching the order in test $D3_P$. The difference of the CORA rating between those models is minor in the regular rating as well as in the alternative rating.

## 7 Discussion

The ratings of the different tests demonstrate the possiblities of an objective rating tool to assess the quality of a dummy model. The most relevant information have to be extracted to define a valid rating procedure of the complete model.

### 7.1 Definition of a model rating procedure

Generally, a dummy model rating procedure should be kept as simple as possible and the results should correlate with experiences of users.

#### *7.1.1 Certification tests*

The assessment of a dummy model by using certification tests seems to be the easiest way to define a rating procedure.

The progress of dummy release 4.5 compared to its predecessor is noticeable but the clearly improved robustness of 4.5 cannot be assessed by CORA. Version 5.0 of the ES-2 model is a big step forward. The ratings of the certification tests are clearly better.

In spite of the good correlation between certification tests and model improvements, this simple procedure is not reliable. It is possible to tune a model to correlate well to the certification test by disregarding the overall performance.

Table 13 shows results of the LSTC ES-2 model (release V0.000.4.ALPHA) exemplarily. Its validation is mainly based on certification tests [7], [8] and its internal geometry is modelled rudimentary. So this model cannot be compared to the FAT ES-2. However, the abdomen of the LSTC ES-2 model achieves a good rating (see Table 13 and Table 1) because of single point optimisation.

Table 13.
Evaluation of certification tests

| | **LSTC** |
|---|---|
| Shoulder | 0.444 |
| Abdomen | 0.784 |
| Pelvis | 0.479 |

In summary, a quality assessment based certification tests might only be helpful to assess the progress of a well-known model but it can fail when using it to benchmark different models of a dummy. Nevertheless, those tests should be part of a rating procedure.

### *7.1.2 Component tests*

Dynamic tests of parts or sub-assemblies might be an important supplement of any assessment procedure. However, they cannot replace tests of the complete dummy. Release 4.5 showed that the rating of the lumbar spine decreased (Table 6) but the overall performance remains almost constant (Table 8).

Furthermore, component tests of all relevant dummy parts and sub-assemblies should be available to define a well-balanced rating procedure based on component tests.

### *7.1.3 Sled tests*

Both sled test series are a solid base of a dummy rating procedure. The number of configurations as well as the wide range of loading conditions is essential to assess models. However, sled tests might not recognise improvements of parts of a model. These minor updates may not relevant in sled tests but might help in vehicle tests. So the rating should be completed by results of component tests.

### *7.1.4 Vehicle tests*

Vehicle tests seem to be the best choice for the evaluation of a dummy model in theory. However, there are strong arguments against the inclusion of those tests in a dummy quality rating.

At first, each vehicle test is unique. There are specific restraint systems, seats and door trims used. So it is very difficult to distinguish between dummy effects and effects caused by the environment. Secondly, the validity of the interior models used is mostly unknown. Thirdly, it is almost impossible to share details of a vehicle simulation with third parties. A dummy model rating procedure requires a well described protocol including all relevant details of the tests used.

A generic test environment (e.g. sled tests) would solve these problems. It reduces the number of unknown or less controlled parameters. So the validity of the corresponding simulations is much higher. Finally, it is easier to publish details of generic tests.

### *7.1.5 Combined rating procedure*

The combination of the three kinds of tests is most likely the best base of a dummy model assessment. A mean rating and a weighted mean rating is shown in

Table 14. The mean rating is using the same weighting factor for each type of tests. Whereas the weighted mean assigns 0.500 to the sled tests and 0.250 to the certification tests as well as to the component tests. The different weighting factors do not change the ranking of the three models.

Release 5.0 was developed by using new test data which covers a wider range of loads. So the CORA rating is clearly the best. Model release 2.0 lost some scores because of the not completed simulation run with the barrier $D4_P$. So the limited numerical robustness is covered by the rating procedure indirectly.

In summary, the combination of certification, component and sled tests seems to be the best approach of a dummy model assessment. However, more component tests should be included into the rating procedure. The chosen weighting factors seem to have only a minor influence on the results.

Table 14.
Evaluation by using various kinds of tests

| | **R2.0** | **R4.5** | **R5.0** |
|---|---|---|---|
| Certification | 0.616 | 0.632 | 0.773 |
| Component | 0.639 | 0.623 | 0.752 |
| Sled (PDB & FAT) | 0.471 | 0.586 | 0.667 |
| Mean | 0.575 | 0.614 | 0.731 |
| Weighted mean | 0.549 | 0.607 | 0.715 |

### 7.2 Build level of auxilary models

The models of the test environment such as pendulum and barriers are identical for the simulations in this study. All results can be compared to each other without any limitation.

These auxiliary models get improved like dummies by the time. It must be analysed and discussed if all simulations of a rating have to use the same auxiliary models. There is significant effort needed to run old dummy models in an updated environment just to update rating results. However, there are first indications that these updates are essential to get valid results. This problem needs further in-depth analysis.

## 8 Conclusion

This study gives a first impression on the possibilities of an objective dummy rating procedure. The rating results of the analysed certification, component, sled and vehicle load cases are reasonable. Furthermore, they mostly correlate to user's experiences. It is the base of the acceptance by users of the model.

The evaluation also shows that a rating procedure must combine different kinds of tests. Certification tests give a limited impression on the overall quality of a dummy model. Component tests can only be used to assess the performance of single parts or sub-assemblies. Sled tests are the right choice for the evaluation of the complete dummy but they might miss improvements of sub-assemblies. Vehicle tests are probably too complex to integrate them into a rating scheme. So finally, a combination of certification, component and sled tests seems to be right mix.

The number of validation tests used in this study is probably too small. Component tests for each relevant body region should be considered. Test data of arm, shoulder, thorax and pelvis is required to evaluate the quality of a side impact dummy.

The influence of the weighting factors of signals and loading cases on the ratings seems to be limited. However, reasonable values must be defined.

Geometry, mass, inertia as well as the application of adequate modelling techniques cannot be assessed by an objective rating tool. So it is essential to check these properties before applying any rating. Otherwise the rating is not valid.

## 9 Limitations

The data used in this study is not sufficient to propose a final procedure. So a larger set of component test should be included in the rating.

Furthermore, it would be helpful to run a full comparison of the FAT ES-2 model and the LTSC model to get a more funded classification on the absolute meaning of the CORA scores.

The responses of head and neck were not analysed in this study. However, a dummy rating should include these body segments.

## 10 References


[1] "ES–2re Side Impact Crash Test Dummy, 50th Percentile Adult Male"; Title 49, Part 572, Subpart U; USA; 2008.

[2] First Technology Safety Systems; "ES-2re, EuroSID-2 50th percentile side impact crash test dummy – User manual"; Part-No. 175-9900; Rev. D 2005-3-10; Plymouth, MI; USA; 2005.

[3] Franz U. et al.; "Observations during validation of side impact dummy models - Consequences for the development of the FAT ES-2 model; 2nd European LS-Dyna User Conference; Gothenburg; Sweden; 2002.

[4] Franz U. et al.; "FAT Side impact dummy models – Remarks on usage and potential pitfalls"; 4th European LS-DYNA User Conference; Ulm; Germany; 2003.

[5] Gehre, C. et al; "Objective rating of signals using test and simulation responses"; 21st International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV); Stuttgart; Germany; 2009; Paper 09-0407.

[6] Gehre, C. et al; "Assessment of dummy models by using objective rating methods"; 22nd International Technical Conference on the Enhanced Safety of Vehicles Conference (ESV); Washington D.C.; USA; 2011; Paper 11-0216.

[7] Maurath, C. et al.; "Overview of LSTC's LS-Dyna anthropometric models"; 11th International LS-Dyna Users Conference; Detroit, MI; USA; 2010.

[8] Mohan, P. et al.; "LSTC/NCAC dummy model development"; 11th International LS-Dyna Users Conference; Dearborn, MI; USA; 2010.

[9] Stahlschmidt, S. et al.; "WorldSID 50th vs. ES-2. A comparison based on simulations."; 11th International LS-Dyna Users Conference; Dearborn, MI; USA; 2010.