

# Metrologically correct decision-making in vehicle testing

Peter, Lau<sup>1</sup> and Håkan, Källgren<sup>2</sup>

*1 MNE-Konsult, Morkulllegatan 1, 42669 Västra Frölunda, Sweden*

*2 Swedish Metrology and Quality AB, Dragspelsgatan 21, 504 72 Borås, Sweden*

**Résumé.** Dans tous les tests les décisions d’approbation sont fondées sur les résultats de mesure qui sont comparés avec les niveaux de tolérance. L’idée météorologique aussi prend en considération l’incertitude de mesure. L’approbation d’un système de frein de voiture aussi signifie que le testeur de frein utilisé conforme à ses spécifications, lequel, en plus de cela, doit être déterminé par une calibration avec une incertitude suffisamment basse. Ainsi nous pouvons discerner plusieurs niveaux d’incertitude -le test d’approbation – la calibration d’équipement – les moyens calibrés ou les références utilisées. Entre ces niveaux un certain proportion de refus, au moins de facteur 3, doit exister, pour que, finalement, une décision correcte (approuvé/pas approuvé) soit possible. Le document explique comment l’incertitude du résultat du test est reliée à la calibration de l’équipement du test, utilisant deux exemples de test de véhicules. Souvent la question, quelle incertitude peut être acceptée dans la calibration d’équipement, est posée par les organisations qui effectuent la calibration et aucune autorité peut déterminer les numéros généraux.

**Key words:** Calibration, testing, traceability, uncertainty, operator reproducibility, test uncertainty ratio

## 1 Introduction

The demand for regularly vehicle testing has existed long time and until quite recently this role was played by national organizations having a monopoly situation. What and how to test was mostly up to the testing organization (Bilprovningen in Sweden, TÜV in Germany etc.). The purpose still is to protect traffic safety and environment. The test programme has expanded over the years due to new functions, new fuels or improved test equipment. Ocular inspection and mechanical control have been complemented with measurement equipment to improve efficiency and trust in assessment.

The overall goal is to keep a certain technical standard and remove traffic dangerous vehicles or those that do not comply with environmental demands. Concerning the measurements performed an answer is requested whether or not a tested function complies with existing tolerances. But it is always the testing engineer that in every instant decides. And he has to state a “clear failure or deficiency”. In order to rely on the measurement results the equipment needs to be traceably calibrated. But only very few strict levels are given concerning metrological demands, which gives space for metrologists to suggest how demands on uncertainty in different stages might look like to get a more neutral and less subjective decision.

One can state that there is an active market for the development of test equipment including various

measurement quantities. Further the monopoly situation has changed to a market of vehicle testing institutions competing with each other where all are accredited (in Sweden by Swedac - Swedish Board for Accreditation and Conformity Assessment, in Norway by Norwegian Accreditation). Before this change a market of service organizations had established providing calibration services for those test equipment. Swedac which is a government authority responsible for quality and safety in Sweden as well accredits these organizations. The regulation authority in Sweden is the Transportstyrelsen (Swedish Transport Agency) that sets up the rules for the testing.

Does this mean we have a regulated area that could be viewed similar to legal metrology? There the legislature sets the demands for measuring equipment to protect consumer interests. And there are accepted rules pushed forward by OIML what calibration laboratories need to achieve with respect to calibration and verification.

## 2 Goal for vehicle testing

Vehicles that do not comply with certain demands need to be repaired or must be removed from traffic;

<sup>a</sup> Corresponding author: plau.gtb@gmail.com

generally the most dangerous or most environmentally unfriendly. This is because a “clear deficiency” must be stated. In contrast to legal metrology rules this is interpreted as in case there is a grey-zone or some uncertainty a tested function cannot be undoubtedly classified as failed. The test should, however, be objective and measurements should serve this purpose as far as possible.

Can one with this perspective claim that low measurement uncertainty helps in decision-making? Of course, if a result is beyond a limit the instrument and method used have less and the object more responsibility for this, which makes a judgement more objective. The subjective “clear deficiency” does not disappear but is easier to state. The task to develop better test equipment, which can be calibrated more accurately, will make pass / no pass decisions easier.

### 3 Measurable quantities for approval/disapproval and demands

All together about 15 different measurements are performed depending on type of fuel, type and age of vehicle and what the inspection is aimed for.

The list contains: Angle and asymmetry of headlights; brake function for foot and hand brake individually on each wheel and possible non-balance between right and left side; retardation; different pressure points; various length, angle and noise measurements (only at special examinations); several gas concentrations in exhaust gas (CO, CO<sub>2</sub>, O<sub>2</sub> and HC) to calculate the  $\lambda$ -value and check optimal combustion – general emission with OBD-system; opacity or light absorption coefficient in diesel fuel exhaust gas; leakage if natural gas is the fuel.

The general demand is that all equipment used for the measurement of a test quantity needs to be regularly calibrated against traceable standards and according to instrument suppliers or the authority responsible for drawing up regulations (in Sweden Transportstyrelsen).

Some of these statements are summarized shortly: Instruments for measuring exhaust gas from diesel engines must have sufficient accuracy. Noise measuring instruments must be of precision type class 2 or better. Length measurement equipment must be within  $\pm 5$  mm at a length of 20 m. Measuring tapes has to follow accuracy class II and callipers follow a standard DIN 862. Timers need to have an accuracy of  $\pm 0,5$  s per minute and angle measurement devices lie within a tolerance of  $\pm 0,5$  °. When calibrating a retardation meter this should be done within  $\pm 3$  %. [1]

The list of real demands is short. The  $\lambda$ -value needs to lie in the interval 0,97 to 1,03. The two belonging gas concentration measurements that contribute to its calculation have a limit between 4,5 vol% and 0,3 vol% CO depending on the type of vehicle, motor and age and 100 ppm and 200 ppm HC depending on vehicle weight. Given an uncertainty for the measurement of the four gases CO, CO<sub>2</sub>, O<sub>2</sub> and HC

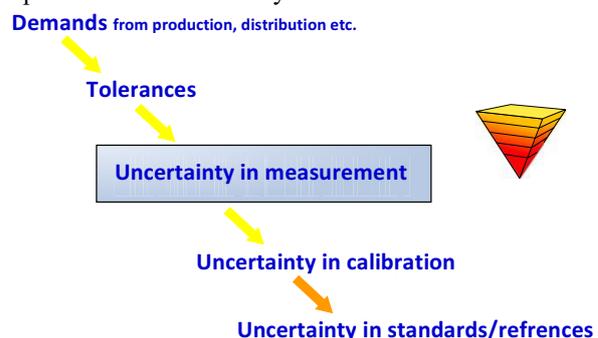
it would be possible to evaluate an uncertainty for the calculated  $\lambda$ -value (Brettschneiders equation). Concerning diesel engines a light absorption coefficient has to lie below  $K=3,5/m$  for turbo engines and below  $K=2,5/m$  for suction engines.

For the brake test no limits are defined which brake force must be achieved. The breaks are operated until almost blockage and the test operator decides approval or not. The only explicit demand concerns a maximum non-balance between left and right wheel pair of not more than 30 %.

What can be stated from this short list is the following. The simpler the measurement the easier is it to quote suitable accuracies for the measurement instrument or the calibration service. Any uncertainty demand chain, however, should start with the measurement uncertainty necessary to decide if a result is inside or outside an allowed tolerance. But if there are no such values it is relative meaningless to express demands for an equipment or its calibration.

### 4 Chain of uncertainty demands in measurement and calibration

A general rule could be the following: In order to judge if a result is inside a tolerance limit the measured value plus the corresponding uncertainty have to be within the tolerance. This would mean the real value, which we never can know, has a probability of at least 95 % to be on the right side. The lower the uncertainty the nearer the result can approach the limit and still be within. If the test uncertainty amounts to maximum 1/3 of the limit value then the measurement uncertainty only accounts for roughly 5 % of the combination of measured value plus uncertainty. Thus one should try to achieve a turn down ratio of at least 3:1 between a limit and the corresponding measurement uncertainty. Figure 1 tries to visualize a demand chain starting from top and down to necessary standards.



**Figure 1.** Uncertainty chain causing demands way down by starting at top.

The same idea also holds for the calibration. The calibration uncertainty should be at least 3 times better than the demanded measurement uncertainty. Further the reference used for calibration should in turn have an uncertainty 3 times lower than the uncertainty aimed for in calibration. This is a rule of thumb before looking into details what makes up the uncertainty budget.

vol% and below.

## 5 Two examples –

### 5.1 Exhaust gas measurement



**Figure 2.** Typical 4-gas meter used at testing stations for calculating the λ -value

As mentioned several acceptance limits exist for carbon oxide in exhaust gas. How the 3:1 thumb rule would be applied to 5 such limits between maximum 4,5 and minimum tolerance 0,3 is demonstrated in table 1 showing what should be expected in terms of uncertainty in test and in calibration with respect to concentration in a reference gas of carbon dioxide.

How does this harmonize with findings in a real measurement starting the other way round, i.e. with the reference gas uncertainty that is given with a 2 %-tolerance of the actual concentration? In the example the gas bottle holds 3,53 vol% CO with an uncertainty of 0,071 vol%. This is an ultimate limit how good a measurement could be and it is clear that there would not be a chance to test levels from 2

**Table 1.** Different CO-limits (vol%) and corresponding 3:1 turn down uncertainty demands.

Maximum permitted CO content	4,5	3,5	2,0	1,0	0,3
Uncertainty in measurement at test	1,5	1,17	0,66	0,33	0,1
Uncertainty in instrument calibration	0,5	0,39	0,22	0,11	0,033
Reference gas CO uncertainty at calibration	0,17	0,13	0,07	0,036	0,011

However, depending on circumstances the uncertainty in calibration can be better than a factor 3. Figure 3 below presents an uncertainty budget for the calibration of a 4-gas meter used at vehicle testing stations.

As can be seen there are seven contributions of which the uncertainty in the reference gas concentration  $u_1$  is the dominating one contributing with over 70 % to the final calibration uncertainty of 0,082 vol%. As the reference gas is the start for the calibration, the calibration uncertainty is the start for the measurement uncertainty (see figure 4).

Model & equation for calibration uncertainty N & normal distribution

$$C_{dev} = C_{read} + \delta C_{rep} + \delta C_{res} - C_{ref} + \delta C_{zero} + \delta C_{moist} + \delta C_{warmup} + \delta C_{leak}$$

S Standard deviation  
R Rectangular distribution  
T Triangular distribution

Uncertainty budget table for calibrating a 4-gas concentration meter using a reference gas CO		ID	value	unit	distribution	standard level	sensitivity coefficient	contribution	rel. importance %	
Xi Input quantities:			xi	δ xi	Φi	u(xi) = δ xi * Φi	ci	ui = u(xi) * ci		
<b>List of influencing parameters</b>										
	Uncertainty in reference gas concentration	δ C <sub>ref</sub>	3,52	0,07	vol%	N	0,0350	1	0,035	73,7
	Instrument reading repeatability	δ C <sub>read</sub>	3,54	0,015	vol%	S	0,015	-1	-0,015	13,5
	Instrument reading resolution	δ C <sub>res</sub>	0	0,005	vol%	R	0,003	1	0,003	0,5
	Adjustment zero point	δ C <sub>zero</sub>	0	0,01	vol%	R	0,006	1	0,006	2,0
	Cleanness and moisture	δ C <sub>moist</sub>	0	0,02	vol%	R	0,012	1	0,012	8,0
u <sub>1</sub>	Possible warm-up shift	δ C <sub>warmup</sub>	0	0,01	vol%	R	0,006	1	0,006	2,0
u <sub>2</sub>	Possible leakage	δ C <sub>leak</sub>	0	0,003	vol%	R	0,002	1	0,002	0,2
u <sub>3</sub>										
u <sub>4</sub>										
u <sub>5</sub>										
u <sub>6</sub>	<b>Output quantity:</b>	<b>C<sub>dev</sub></b>	<b>0,02</b>	<b>combined standard uncertainty U<sub>c</sub></b>			<b>k=1</b>	<b>0,041</b>	vol%	
u <sub>7</sub>				<b>expanded uncertainty U<sub>k=2</sub></b>				<b>0,082</b>		2,32%

**Figure 3.** Uncertainty budget for calibrating a 4-gas meter with CO at a level of 3,52 vol %.

Model equation for measurement uncertainty

$$C_{meas} = C_{read} (-C_{dev}) + \delta C_{cal} + \delta C_{reprod} + \delta C_{drift} + \Delta C_{dev}$$

Uncertainty budget table for measuring &		ID	value	distribution	unit	frequency	standard level	sensitivity coefficient	contribution	rel. importance %	
CO <sub>2</sub> content in exhaust gas		X <sub>i</sub>	x <sub>i</sub>	δ x <sub>i</sub>	φ <sub>i</sub>	u(x <sub>i</sub> ) = δ x <sub>i</sub> * φ <sub>i</sub>	c <sub>i</sub>	u <sub>i</sub> = u(x <sub>i</sub> ) * c <sub>i</sub>			
<b>List of influencing parameters</b>											
Calibration uncertainty	δ C <sub>cal</sub>	3,72	<b>0,082</b>	vol%	N	0,0410	1	0,041		50,8	
Reproducibility in using instrument	δ C <sub>reprod</sub>	0	<b>0,020</b>	vol%	S	0,020	1	0,020		12,1	
Possible drift of instrument since calibration	δ C <sub>drift</sub>	0	<b>0,010</b>	vol%	R	0,006	1	0,006		1,0	
Probably not corrected deviation	Δ C <sub>dev</sub>	0	<b>0,020</b>	vol%	R	0,012	1	0,012		20,1	
<b>Output quantity:</b>		<b>C<sub>meas</sub></b>	<b>3,72</b>	<b>combined standard uncertainty U<sub>c</sub></b>			<b>k=1</b>		<b>0,058</b>		vol%
		<b>δ C<sub>meas</sub></b>		<b>expanded uncertainty U<sub>k=2</sub></b>					<b>0,12</b>		3,09%

Figure 4. Uncertainty budget for measuring exhaust gas CO with a 4-gas meter testing a vehicle.

Here four contributions are important. First is  $u_1$  coming from the calibration amounting to about 50 % of the testing uncertainty. Second important is the last one  $u_4$ . This contribution corresponds to the detected but not adjusted error in the 4-gas meter. As it is likely that the correction of -0,02 vol% is not performed in testing, this systematic error is added to the measurement uncertainty in a linear way (maxi-

mum combination). This contribution could largely be avoided if the instrument were adjusted during calibration or if a corresponding correction were applied in testing, which most probably is not the case. Thus the total testing uncertainty actually could be reduced significantly. A picture how this finding matches the factor 3:1 rule of thumb is shown in figure 5.

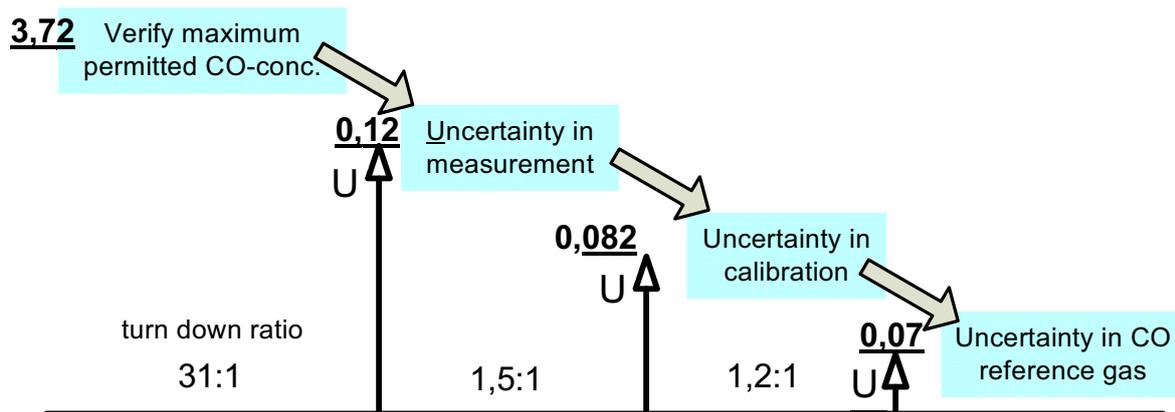


Figure 5. Example for a realistic measurement and preceding calibration uncertainty.

Compared to the actually measured CO concentration a turndown ratio of 31:1 is reached

in the measurement step, whereas the preceding steps to the right have only half of the 3:1 rule of thumb turn down ratio. In this example actually it is possible to verify all levels except the lowest one of 0,3 vol%. To

have this picture on the retina can be useful when demanding a calibration service to improve its calibration uncertainty or a testing facility to improve measurement uncertainty. Everything is depending on the uncertainty level of the reference gas concentration. Authorities not being aware of this hierarchically uncertainty chain might actually demand the impossible.

For HC-measurements with the two levels of 100 and 200 ppm the same chain could be demonstrated.

### 5.2 Brake force measurement

In brake testing the only defined demand is a maximum 30 % non-balance between measured left and right wheel brake force. A typical brake force is at least 2000 N and a 30 % non-balance corresponds to 600 N giving for example 1700 N for the left and 2300 N for the right wheel brake. One third would demand an uncertainty of 200 N at maximum for the difference measurement.



Figure 6. Roller brake tester for a wheel pair

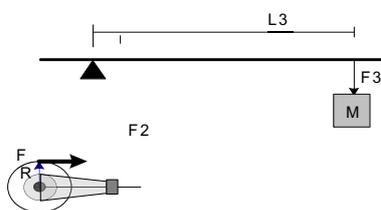


Figure 7. Lever system to apply a reference force in roller brake tester calibration

Figure 6 shows a typical roller brake tester and figure 7 how a retarding reference force is

applied to the motor axis using weights and a lever system. Using different weights or various lever arm length one can simulate different discrete brake forces. This is the basis to define the measurement model for the calculation of the calibration uncertainty.

**30% at 2 kN  
= 600 N**

**Maximum  
permitted non-  
balance**

It should be mentioned that sometimes also continuous brake forces are applied using calibrated load cells as reference and totally new brake test equipment is currently developed.

Assuming an experimentally determined reproducibility for the difference measurement of 150 N at a test station and a reasonable drift since last calibration of 60 N both for the right and left roller pair one would fulfil a 3:1 turn down ratio if each side could be calibrated with an uncertainty of maximum 72 N (see fig 8).

A reasonable calibration uncertainty for a 2 kN brake force is expect to lie between 30 and 50 N. This means the demand of  $U = 72$  N during calibration of each side should be achievable. If a nonbalance of less than 600 N is detected the actual uncertainty is not questioned.

In the calculation above no correlation is assumed. In reality it is reasonable to assume a drift in the same direction for the right and left roller pair. And several uncertainty contributions in calibration are highly correlated reducing the uncertainty in the difference measurement. Among those one can mention the weight and length of all arms in the lever system (contributions  $u_3$  to  $u_7$  in the table below).

Thus, if the brake force at average is about 2000 N the testing uncertainty is acceptable. At higher brake forces the calibration uncertainty will increase steadily but so will of course also the uncertainty corresponding to a factor of 1/3 of the difference.

This means the calibration accuracy will most probably be sufficient the higher the brake force. However, the test station needs to show that the reproducibility when testing vehicle brakes for nonbalance between right and left is under control i.e. all personnel needs to perform the job the same way. Reproducibility here means different test persons, different cars and several repetitions in each case.

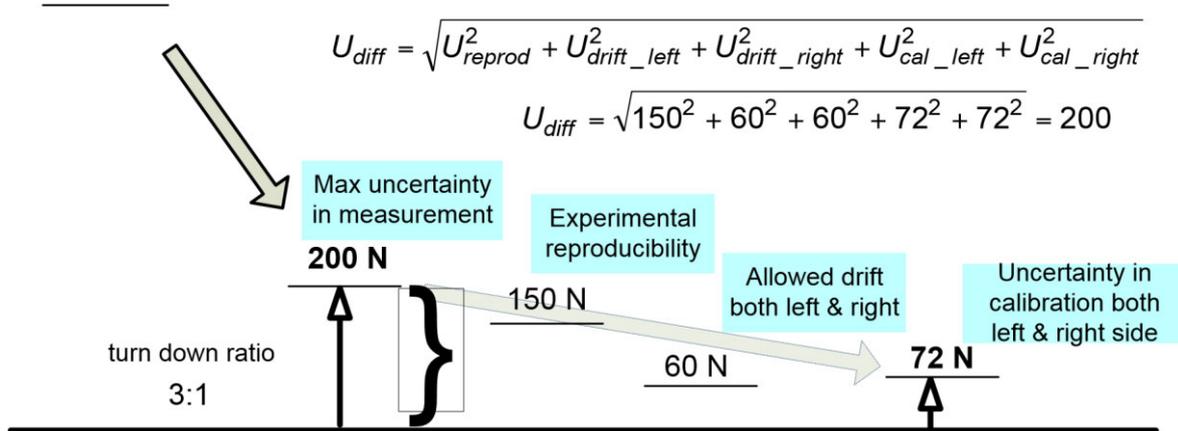


Figure 8. Scheme for estimating a demand for calibration uncertainty.

Model equation for calibration uncertainty

$$F_{dev} = F_{read} + \delta F_{rep} + \delta F_{res} - \frac{(L_2 + \delta L_2) \cdot (L_3 + \delta L_3) \cdot \cos(\alpha + \delta \alpha) \cdot (M + \delta M) \cdot (g + \delta g)}{(L_4 + \delta L_4) \cdot (R + \delta R)}$$

N normal distribution  
 S Standard deviation  
 R Rectangular distribution  
 T Triangular distribution

Uncertainty budget table for calibrating roler brake tester - one side										rel. impor- tance %	
Xi	ID	value	distrib- ution	unit	fördel- ning	standard- level	sensitivity coefficient	contri- bution			
Input quantities:		xi	δxi		ϕi	u(xi) =δxi*ϕi	ci	ui =u(xi)*ci			
<b>List of influencing parameters</b>											
u <sub>1</sub>	Average reading	F <sub>read</sub> / δC <sub>rep</sub>	2040	<b>8,22</b>	N	S	8,2199	1	8,220	17,8	
u <sub>2</sub>	Instr. Resolution	δC <sub>res</sub>		<b>0,50</b>	N	R	0,289	-1	-0,289	0,0	
u <sub>3</sub>	Radius R_of roller	R / δR	108	<b>1,0</b>	mm	R	0,577	-19	-10,696	30,1	
u <sub>4</sub>	Length of force arm L_2	L_2 / δL_2	178	<b>1,0</b>	mm	R	0,577	11	6,492	11,1	
u <sub>5</sub>	Length of force arm L_3	L_3 / δL_3	773	<b>1,0</b>	mm	R	0,577	3	1,494	0,6	
u <sub>6</sub>	Length of force arm L_4	L_4 / δL_4	50	<b>0,5</b>	mm	R	0,289	-40	-11,551	35,1	
u <sub>7</sub>	Mass of applied weight M	M / δM	8,00	<b>0,005</b>	kg	R	0,003	250	0,722	0,1	
u <sub>8</sub>	Possible error in angle a	α / δα	0	<b>5</b>	°	R	2,887	2	4,396	5,1	
u <sub>9</sub>	Earth acceleration g	g / δg	9,82	<b>0,003</b>	m/s <sup>2</sup>	R	0,002	204	0,353	0,0	
	Reference force applied		2000,7		N						
<b>Output quantity:</b>		<b>F<sub>dev</sub></b>	<b>39,3</b>	<b>combined standard uncertainty U<sub>c</sub></b>			<b>k=1</b>		<b>19,5</b>		
				<b>expanded uncertainty U</b>			<b>k=2</b>		<b>39</b>		1,91%

Figure 9. An uncertainty budget valid for the calibration of a typical roller brake tester using weights

demonstrate how a maximum acceptable uncertainty in measurement and consequently in calibration could be derived.

## 6 Summary

The area of vehicle testing is not regulated in a strict manner. Whereas there are detailed demands for a few measurements performed during vehicle testing the majority of these have no demands at all (no tolerances, no accuracy concerning the equipment used or the uncertainty in testing that should be achieved). Even for the few test limits pointed out there are no demands for a measurement uncertainty, which is due to the variation in type and age of vehicle. This generates difficulties to define limits and prescribe suitable demands for the related measurement uncertainty during vehicle inspection. If, however, such demands existed the two examples can

## 7 Conclusions

Legal metrology has largely succeeded to internationally harmonize demands on the measurement quality when using certain equipment and thus generated trust that consumers can have confidence in mass or volume information in purchased products. A similar process regarding vehicle testing with its various test and decisions based on measurement results has only recently started in connection with the opening of an earlier monopolized market. Within the frame of legal metrology OIML has played an important role and national requirements

have adopted the international demands. Within the frame of vehicle testing there are no such rules. There the suppliers of test equipment largely set goals mainly due to marketing concurrence what quality in testing can be achieved.

Prescribing authorities (Transportsyrelsen in Sweden) probably lack competence and currently only point out that all equipment used in testing needs to be calibrated. Accreditation bodies (in Sweden Swedac) regulating this area since the end of the monopoly situation try to imprint by inspecting the accredited service organizations, but cannot set up goals for which uncertainties that are acceptable in testing and which are not. However, even if it is possible to construct reasonable uncertainty chains for testing and calibration from a pure metrological point of view there are two circumstances that prevent such a development. The one is the lack of clear limits for go-no go borders in testing, which is intimately linked to the goal to not approve only at clearly stated deficiencies. Second a well-educated and motivated test personnel probably has better qualification to live up to this goal despite subjective judgements of what a “clear deficiency is than a strictly regulated demarcation.

## References

- [1] **TSFS 2015:18** Föreskrifter om ändring i Transportstyrelsens föreskrifter (TSFS 2010:78) om teknisk kontrollutrustning hos besiktningsorgan