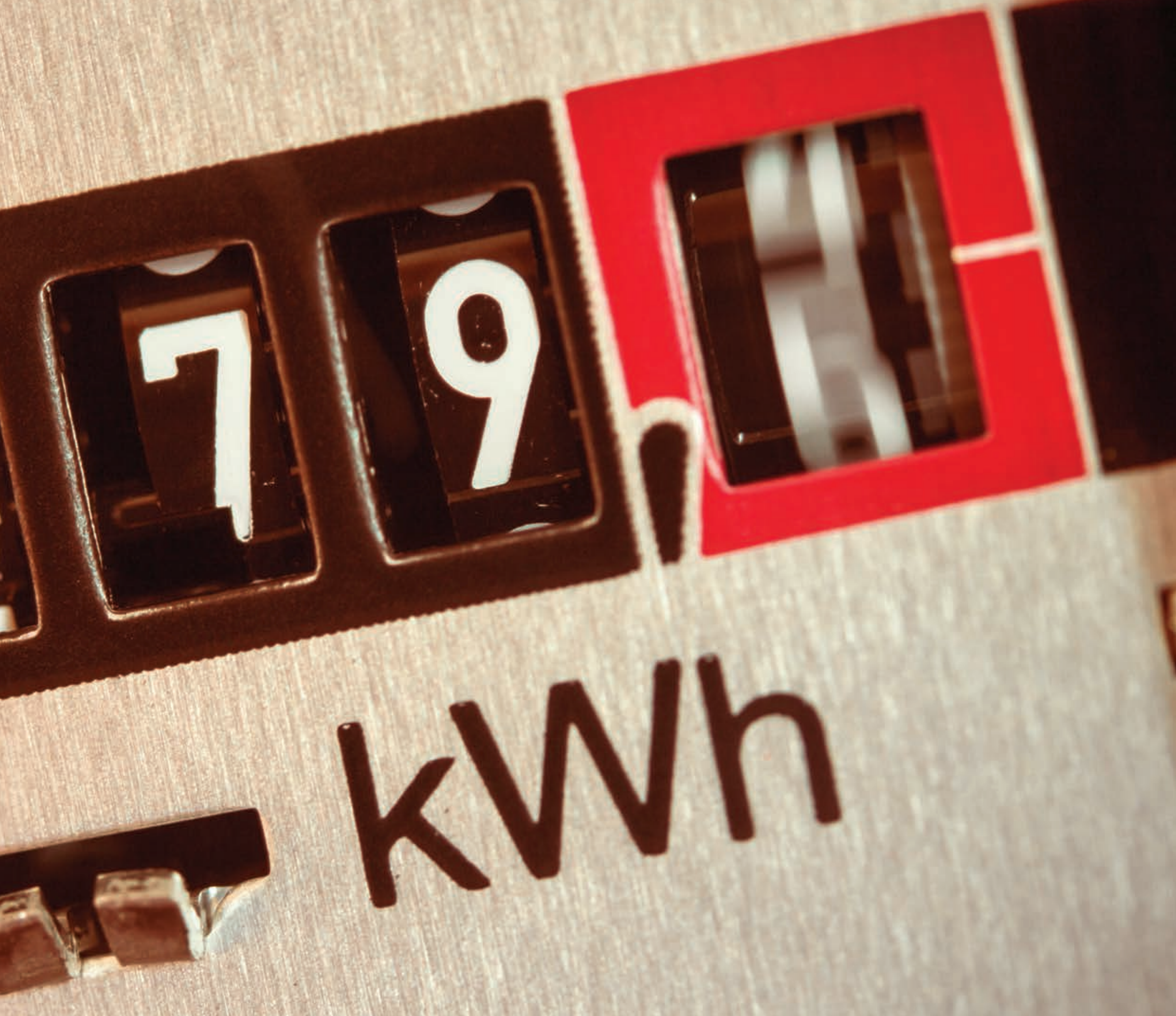
CLOSING THE 'REALITY GAP' – ENSURING A FAIR ENERGY LABEL FOR CONSUMERS

IDENTIFYING WEAKNESSES AND RECOMMENDING SOLUTIONS TO IMPROVE CRITICAL ASPECTS OF TEST STANDARDS FOR TELEVISIONS, REFRIGERATORS AND DISHWASHERS



Authors

This report is a collective undertaking by four different organisations: CLASP, ECOS, EEB and Topten.

The contributing authors are respectively:

www.clasp.ngo

Marie Baton
Mike Scholand

www.ecostandard.org

Justin Wilkes
Christoforos Spiliotopoulos

www.eeb.org

Stephane Arditi
Jack Hunter

www.topten.eu

Anette Michel
Helene Rochat
Eric Bush

ECOS led on the selection of the products to study. All STEP partners contributed to the design of the test plans to explore issues in the measurement standards. CLASP led on the testing of televisions and fridge-freezers and Topten led on the testing of dishwashers. EEB led on the communication of the findings and recommendations to standardisation bodies, the Commission and other stakeholders.

The authors would like to express their gratitude to the European Climate Foundation (ECF) and the ClimateWorks Foundation for supporting STEP. They would also like to thank the French Agency for Energy and the Environment (ADEME) and the Swiss Federal Office of Energy for additional funding support for specific product testing. Finally, the authors gratefully acknowledge Francisco Zuloaga from ECF for his review of this report.

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

French Environment & Energy Management Agency

www.europeanclimate.org **www.bfe.admin.ch** **www.ademe.fr** **www.climateworks.org**

21 June 2017

Note: the tests performed for this report do not follow the official market surveillance procedure and sometimes deviate from the harmonised European and/or international test standards. For that reason, the test results shown in this report should not be construed as market surveillance tests and have no bearing on legal compliance.

With the support of
the LIFE Programme
of the European Union

This communication reflects the authors' views and does not commit the donors.

Designed by www.dougdawson.co.uk

EXECUTIVE SUMMARY

The A-G energy label for TVs, fridges and a host of other appliances is perhaps the best-known symbol of Europe after the Euro currency symbol. That privilege may be justified if you consider that energy efficiency policies for products are set to reduce the average home energy bill by nearly €500 per year by 2020. In reality, such savings will only be realised if consumers trust the label and continue buying energy efficient products.

The STEP project was designed to help ensure the on-going trust of consumers in the policy measures that promote energy-efficiency across Europe. The STEP project investigated the test standards of three product groups – televisions, freezers / fridge-freezers, and dishwashers sold in Europe. The Energy Label relies on European harmonised test standards (EHTS) for efficiency measurements. Energy labelling regulations, defining class thresholds and what information should appear on the label, are policies defined by the European Commission and Member States, whereas EHTS are measurement methods and procedures, developed by the European standardisation body CENELEC (for the three products considered), under a mandate from the European Commission[1]. We conducted both EHTS tests and deviations from these tests, to explore the performance of these products both in standardised conditions and in conditions that are more representative of real life. The investigation was limited, with only one unit of each model tested, while official market surveillance tests use larger samples.

The STEP findings suggest four principal concerns with EHTS that may be undermining the accuracy of the label: (1) differences in energy consumption between EHTS and ones that reflect real world usage; (2) EHTS that do not keep pace with technological progress; (3) ambiguities in EHTS that undermine reproducibility of parameters measured; and (4) confusing or non-existent consumer information.

1. In some cases CENELEC adopts, with or without modifications, pre-existing standards developed at the international level by the IEC (International Electrotechnical Commission). They then publish an EN (European Norm) version of an IEC standard. The European Commission can also directly refer to an IEC standard.

KEY FINDINGS

1. Differences in energy consumption between EHTS and ones that reflect real world usage

Standardised tests used to measure products do not always reflect real life conditions. Certain standardisation bodies have favoured repeatability and reproducibility of tests in the laboratory over being reflective of realistic consumer use. We observed this difference in our measurements for all three product groups studied. For example, dishwashers are only tested on a very efficient, but infrequently used wash programme. TVs are tested with a video clip from 2007 that does not reflect typical home viewing or increasingly common TV technologies Tests with a modern video clip in a higher quality format saw one model consume double the power. Fridge-freezers are tested without opening the doors and without any load in the fresh food compartments. When we conducted door opening tests, we found it caused four models to consume significantly more energy, with one model continuing this higher level of energy consumption for over 24 hours. Adding test samples representative of food inside the fridges caused higher energy consumption in all three models tested. Beyond obscuring real-life energy use, unrealistic EHTS also make detecting test conditions easier, thus increasing the risk of circumvention. Some cases where differences were striking between declared performance using official test methods and measurement by the STEP team have been reported to market surveillance authorities for further investigation.

2. Test standards that do not keep pace with technological progress

Certain EHTS we investigated were found not to be fully suitable for testing all aspects of products placed on the market today. This issue was particularly true for televisions, whose standard test procedure (EN 62087:2016[2]) contains a test video clip designed 10 years ago and still being used today. With internet connectivity becoming a default feature, the current EHTS is unable to capture all aspects of the new televisions. Software updates, for example, caused energy consumption to rise by about a third in three TVs when tested with a new ultra-high definition (UHD) video test clip. The standard video clip does not measure high dynamic range (HDR), which was found to use more energy in some televisions tested, but is not part of today's standard test. We observed that automatic brightness control (ABC) can cut power use by between 32% and 76%, but there is no EHTS to measure it. These deficiencies in EHTS can lead to inaccurate performance declarations and unreliable labels for consumers. In one extreme case, the difference in power consumption between an UHD-HDR format and a high definition (HD) format was as much as 130%.

3. Ambiguities in EHTS undermine reproducibility of test measurements

Standardisation bodies work hard to make test standards as clear and concise as possible. However in some EHTS there can be ambiguities around how to apply certain measurement procedures which introduces an undesirable element of variability in the measurement.. An example of this was found with fridge-freezers and the measurement of interior volume. The current EHTS does not make it clear enough how to measure the volume, despite this value being a direct input to the calculation of energy label class. As a consequence, possible differences between declared performance and measured performance can be exacerbated. The STEP team found some discrepancy in 6 out of the 10 appliances tested when comparing declared and measured storage volumes.

4. Confusing or non-existent consumer information

Ecodesign and energy labelling regulations, and subsequently EHTS, do not require consumer information on all aspects of features that could affect the energy consumption of products. For example, when changing the default picture settings of televisions, we found that energy saving features were deactivated in five models without informing the consumer. For two of those five models, the energy saving features were greyed-out and could not be re-enabled without a factory reset – despite the fact that these features contributed to the energy label class rating. Energy labels for dishwashers are based on the Eco programme. It is rarely used, perhaps because it is one amongst the many proposed programmes (including the programme typically called ,'normal'). Adding wash functions on top of the Eco programme can also increase energy use by 30-50%. There is no requirement for providing that information on these and similar aspects. Freezers and fridge-freezers suffer from confusing performance controls and settings. In STEP testing, we found that lowering the internal temperature by 1 degree consumed 4-8% more energy, but no requirement exists to warn users of the impact or help them manage the temperature settings. Modes that would imply energy savings were found to produce no savings compared to measurement made in standard conditions and one model actually consumed 50% more energy in that mode.

2. EN 62087 :2016: https://infostore.saiglobal.com/store/details.aspx?ProductID=1841650

Policy recommendations

- Accommodate real-life conditions by better reflecting product and consumer behaviour in test standards. Notably, this should be explored during ongoing revisions of ecodesign and energy labelling regulations for dishwashers, televisions and fridge-freezers as well as through the appropriate CEN/CENELEC and ISO/IEC standardisation bodies. This key principle should be an integral part of all product regulations and test standards. An up-to-date video loop should be used when testing televisions, fridge-freezer doors should be opened more systematically and dishwashers should be tested on a more comprehensive set of programmes. These recommendations echo provisions in the newly revised Energy Labelling Directive.

- Update EHTS more frequently to reflect market and technological developments, particularly for products with rapid technology evolution and/or with new features added. If possible, incremental improvements should be made to existing standards to speed up the integration of new functions. In the case of televisions however, we call for a replacement of the video test loop, as today's measurement standard was created in 2007 and is obsolete.

- Require that consumers are given helpful information on energy impacts when they change settings. For connected appliances, market surveillance testing should be conducted after the software has been updated, if an update is available.

- Base energy label classes on the normal/most commonly used programmes or modes, not only the most efficient but potentially infrequently used eco and energy-saving modes. Consumers should be offered information on when to use these saving modes and the energy savings they can expect.

- Expand and improve market surveillance and enforcement. Today, cooperation between national enforcement authorities is optional, despite the fact that EU product policies are clearly structured around this single market instrument. Participation in the Ecodesign and Energy Labelling Administrative Cooperative (ADCO) should be mandatory, and funding should be provided to ensure more systematic coordination and sharing of test results at the European level. Negative publicity can be more effective than fines, and thus enforcement authorities should consider publicising grievous cases of non-compliance and test standard circumvention. We support the principle put forward by the European Parliament, that consumers should get compensation for products that consume more energy than was declared by manufacturers, even when the situation was identified outside of the legal warranty period.

- Supplement EHTS within an additional test defined within real-world boundaries. approach is being considered by the automotive industry to prevent EHTS circumvention: EHTS tests are followed by randomised tests in the real world, under driving conditions that fit within defined boundaries. This could be applied for energy related products. If the difference between the EHTS and randomised tests is within acceptable limits, the product is considered compliant. If it is outside the acceptable limits, the model is studied further and/or declared non-compliant. Such tests would become part of the market surveillance procedures but not be replacements for the EHTS, which itself should continue to be improved.

Standardised tests used to measure products do not always reflect real life conditions. Certain standardisation bodies have favoured repeatability and reproducibility of tests in the laboratory over being reflective of realistic consumer use.

CONTENTS

INTRODUCTION: WHAT IS THE SMART TESTING OF ENERGY PRODUCTS (STEP) PROJECT?

CONTEXT OF STEP

Energy efficiency standards and labels (EES&L) are one of the most cost-effective ways to reduce carbon emissions through market policies and programmes. Through a mixture of policy measures that push and pull markets[3], they have a major impact on the choices consumers make when purchasing energy consuming appliances and products. Thanks to the EU Ecodesign Directive and Energy Labelling Directive, the average product in Europe will, by 2020, do the same job using around one fifth less energy[4]. The resulting primary energy savings from these Directives amount to 9% of the total EU energy consumption, contributing to savings of nearly €500 per year on energy bills[5] for the average household. Most Europeans are familiar with European EES&L thanks to the colourful A-G energy label found on all televisions, fridges, vacuum cleaners and other domestic and commercial appliances being sold.

Delivering even more energy savings for Europe requires EES&L that are well designed and accompanied by robust, appropriate test standards. EES&L regulations and test standards are interrelated and influence each other. But the process of setting a test standard after a regulation has been adopted is slow, and may lead to tests becoming irrelevant through market and technological evolution, as highlighted in a 2014 report[6]. Another key dimension is proper enforcement of regulations. In Europe, research shows that compliance can be improved across countries and products, with low compliance for some types of appliances[7]. Overall it is estimated that 10% of total energy savings is lost due to non-compliance[8], equating to just over €10 billion in higher energy bills each year.

The STEP project analysed whether test standards reasonably reflected real life usage of products – what can be called the *representativeness* of test standards. An iconic example are vacuum cleaner energy labels that are based on a test made with a completely empty bag, while often in real-life use in the home they can be part-loaded and consuming more energy. The STEP partners are concerned that if lab tests reproduce conditions that differ fundamentally from real life conditions, labels would lose their credibility as a guide to expected energy performance.

Furthermore, as appliances and products become increasingly sophisticated and 'smart', they may also become better able to detect specific test conditions set out in the EHTS and adjust their performance and energy consumption. If a test is very similar to real life, it becomes more difficult for software to differentiate between test conditions and real life. By assessing the representativeness of test standards we also indirectly assess the risks of test standard detection and circumvention. Investigating where EHTS could be prone to circumvention was also an important part of this project. To explore this, we conducted tests with minor changes compared to EHTS and studied the energy impact. If energy use was significantly higher than expected, it could imply the machine has been able to recognise the EHTS and adjust its energy consumption to achieve a more energy-efficient rating. We were trying to assess the gap between consumption using EHTS and expected consumption using slightly different tests methods aiming at reflecting better real life conditions.

More broadly, the STEP team evaluated possible shortfalls of testing standards, and their ability (or not) to capture energy performance and related savings potentials stemming from rapid technological development. Also, as more connected, smarter appliances enter the market, the differences between the test and how equipment are used are likely to become more widespread and difficult to address. For example, products that receive software updates over the Internet could change their energy settings through an update and end up increasing energy consumption without the consumers' knowledge and agreement.

In this context, the STEP project was launched to investigate whether and how these issues are present in the products and equipment covered under EU Ecodesign Directive and Energy Labelling Directive policies.

3. See Coolproducts: https://www.coolproducts.eu/products-are-changing
4. European Commission: https://ec.europa.eu/energy/en/news/report-eu-energy-efficiency-requirements-products-generate-financial-and-energy-savings
5. European Commission: https://ec.europa.eu/energy/sites/ener/files/documents/ecodesign_factsheet.pdf
6. Developing Measurement Methods for EU Ecodesign and Energy Labelling Measures, a discussion paper. Edouard Toulouse, published by CLASP Europe. February 2014. http://clasp.ngo/Resources/Resources/PublicationLibrary/2014/Alignment-of-EU-Test-Procedures-and-SL-Regulations
7. See table on pages 9-10: http://www.energylabelevaluation.eu/tmce/Literature_report_Energy_Labelling_Ecodesign
8. See: http://www.web4948.vs.speednames.com/upl/File/Ecodesign/Session-2-CLASP.pdf

AIMS

The STEP project aims at identifying opportunities for capturing even greater CO_2 emission reductions in Europe from product standards and labelling through:

- Identifying issues and potential failures in standardised product testing;
- Documenting discrepancies between real-life and declared performance; and
- Suggesting improvements to the ecodesign and energy label community and the standardisation community to address any discrepancies or declaration problems which are found.

DESCRIPTION OF STEP

STEP is a project team, consisting of four non-profit, non-governmental organisations (NGOs): CLASP Europe, the European Citizen's Organisation for Standardisation (ECOS), the European Environment Bureau (EEB), and Topten International. STEP was funded by the European Climate Foundation and the ClimateWorks Foundation with additional support for specific product testing from the French Agency for Energy and the Environment (ADEME) and the Swiss Federal Office of Energy.

Phase 1 Screening and prioritising: Screen all product groups covered under ecodesign and energy labelling policy and identify those for which the standard tests are known to suffer poor representativeness compared to real life usage. A final prioritisation of products for lab testing included new opportunities linked to the revision of certain product ecodesign and energy labelling implementing measures. The three product groups selected were: televisions, fridges/ freezers and dishwashers.

Phase 2 Testing and interpreting: Purchase products selected in Phase 1 and test them using both the EHTS as well as a deviation from the standard test, for example by checking more 'real life' usage. All the tests have been performed in independent laboratories certified to test products on behalf of national market surveillance activities or with extensive experience doing tests and research on the products selected.

Phase 3 Communicating results: Raise awareness about the findings to the European Commission, market surveillance authorities, standardisation bodies, members of the European Parliament and the general public as appropriate.

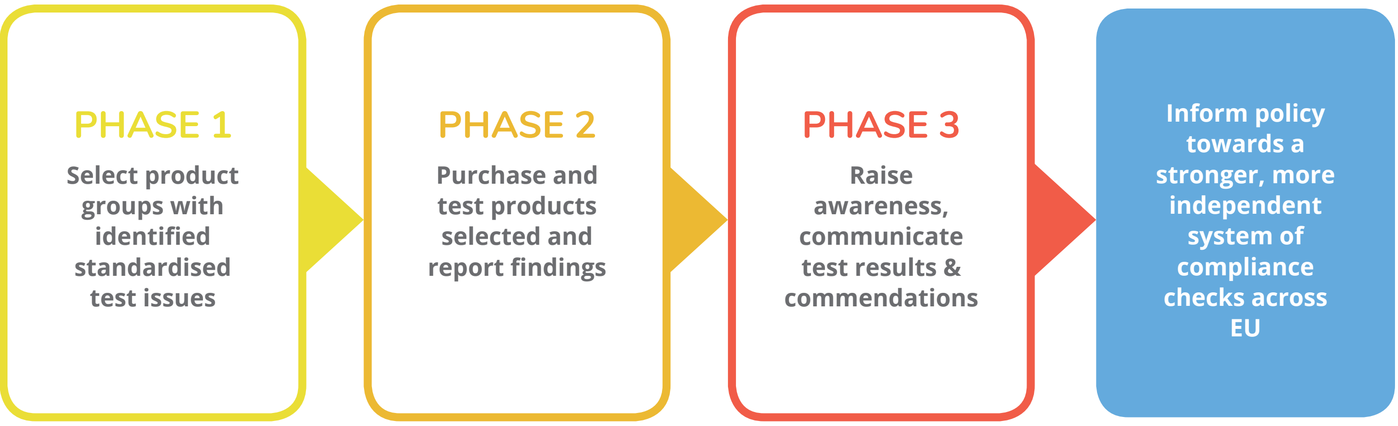


Figure 1. Flow diagram of the three phases of the STEP project

Twenty products were tested, including 7 televisions, 10 fridges/freezers and 3 dishwashers. The sample sizes were limited to one model due to time and budget constraints. However, the range of models strived to find a balance between variety of appliances on offer within each product group and the latest market trends with regard to technology. As we aim to inform future development of standardisation policy, we selected models that were either high end or high-volume, over a wide range of prices. We also selected models that had high energy efficiency classes that are good indicators of future features and technologies for mainstream products. In Annex I of this report, we provide an outline of the testing steps and procedures that were followed for the three product groups tested.

LIMITS OF TESTING AND INTERPRETATION

The test results referred to in this report were only conducted on one unit of each model. Thus, they are indicative findings, rather than conclusive findings from a market surveillance point of view. The official verification procedure, as defined in the European regulations, may require the testing of several products or appliances of the same model to ensure the measured results are robust enough. Thus, our test results cannot be considered proof of regulatory compliance or non-compliance. For this reason, brands and model numbers of the products tested are not included in this report.

Where our test results found a unit to be less efficient than declared, or found a larger than expected difference between the EHTS test and one where we had deviated from the standard, we informed market surveillance authorities. The purpose of this project was not to assess regulatory compliance, but rather to identify shortcomings in EHTS and suggest appropriate remedial actions.

STRUCTURE OF THIS REPORT

This report builds on tests performed on televisions fridges/ freezers and dishwashers. A separate report was issued on the dishwasher testing in May 2017, as this product group had received co-funding for the testing from the Swiss Federal Office of Energy (SFOE) who required the publication of the test findings. Thus, the dishwasher report can be found on the Topten Europe website[9]. No suspicion of non-compliance or test circumvention was detected for any of the three dishwasher models tested.

The STEP partners identified four overarching problems relating to test standards which were derived from our study of the three product groups:

- **A lack of representativeness of test standards** – more realistic test methods could result in more accurate quantification of performance, more accurate labelling and, potentially, more accurate energy savings potential assessment;
- **A lack of suitability of test standards** – certain characteristics of product performance are not measured effectively or simply do not have a test method;
- **An ambiguity of test standards** – existing standards are not precise enough or less adequate given how product technologies have changed; and
- **A lack of user information** – product modes can vary and the impact on energy consumption is not properly or clearly communicated.

These four main issues will be documented in the first part of this report using examples of findings picked from individual test work. Part one will conclude with a reference to a few cases which might merit further analysis by market surveillance authorities. In the second part, some recommendations will be made to help improve test standards and an innovative testing approach will be suggested.

9. Click here: http://www.topten.eu/uploads/File/Topten_Dishwashers_May17.pdf

PART I: TEST STANDARDS: ARE THEY FULFILLING THEIR ROLE?

The role of test standards for products and appliances

Test standards define the methodologies used by manufacturers to declare the performance of their products, and by market surveillance authorities and independent laboratories when verifying the regulatory compliance of products placed on the EU market. They are used for the assessment of minimum performance requirements to be met to place a product on the EU market as defined under the Ecodesign Directive, as well as for the determination of the energy label class, the measurement of energy consumption and all other parameters referred to in the energy label of products. They define the step-by-step process to follow and the details of each step with regards to the type, the conditions and the procedure of measurement. For example, to verify the energy performance of fridges/freezers, the room temperature to perform the test should be normalised at 25°C in the EHTS for fridge/freezers. For dishwashers, the colour, type and even the soiling of the dishes is prescribed.

Dishes soiled and waiting for washing under lab conditions

For EHTS to fulfil the role described above, several criteria have to be considered. Firstly, having a precise and repeatable methodology is essential to compare products with a given baseline and any ambiguity or open interpretation of the test method should be minimised. This is crucial to maintain a fair assessment and comparison of the performance of products and to ensure a level of reproducibility of the test in different manufacturing sites and laboratories around the EU and the world. However, consideration of repeatability and reproducibility alone do not constitute an appropriate test method[10]. A lack of adequate consideration of the representativeness of real-life conditions can lead to results that are irrelevant to consumers. The dominance of test measurement repeatability and strict test conditions can sometimes be counter-productive, decoupling the measured quantities from real life use of the product. There is a risk of standards becoming out-dated with regards to contemporary usage patterns and obsolete with regards to technological progress. Oversimplification and artificiality of test procedures can also make them vulnerable to circumvention.

The television testing standard, EN 62087:2016[11], was revised in 2016 without major changes to the methodology or the test video clip that underpins the power measurement. This was identified as a concern by the STEP team because the video test clip is not representative of typical viewing content and it isn't available in ultra-high definition (UHD) resolution or high dynamic range (HDR). Furthermore, the standard does not offer a method for measuring the performance of automatic brightness control (ABC), a recent energy saving feature that dims the screen when the ambient light levels are lower. All three features are becoming mainstream and can have energy use implications.

The European harmonised standard for fridge-freezer testing is EN 62552:2013[12]. Standard performance testing of refrigerators and freezers has been redefined at the international level in IEC 62552:2015[13], which includes significant differences compared to EN 62552:2013. IEC 62552:2015 is being proposed by CENELEC to the Commission as a basis for the preparation of a future European harmonised standard for refrigerators and freezers, expected to be adopted in 2018. One of the aims of this new standard was to correct various shortfalls and uncertainties linked to EN 62552:2013. The basic energy consumption test is, however, still realised in steady state operation (i.e., stable ambient temperature, humidity and settings, no door openings, etc.) and with empty compartments. Freezer compartments used to be filled with test packs in EN 62552:2013, but are tested empty in the new IEC version. This is of course very different from real life conditions where air flux varies too much to allow the appliances to reach a perfectly steady state. The STEP team designed tests to assess whether these could represent significant issues for energy consumption declarations and compared the performance of several models.

The dishwasher testing standard is defined in EN 50242:2008[14]. It is noted that this standard has been updated to correct some shortcomings already. In 2016, CENELEC revised the EN standard, based on mandate M481[15], and at the same time aligned it to changes that have been made in the IEC standard (IEC 60436:2015)[16]. The mandate included a task: *"To ensure that the prospective harmonised standard(s) includes a procedure that avoids an appliance being programmed to recognise the test cycles, and reacting specifically to them."* (EC, 2012)

Notably, a more integrated approach was adopted to test different functionalities of the dishwashers. By combining tests for cleaning and drying functions, for example, the updated standard intends to better approximate real life usage. In daily life, people do not run different cycles for cleaning and drying. This combined testing reduces the 'artificiality' of testing, consequently reducing the risk of test conditions detection and circumvention. In the updated standard, the test load better reflects consumer use by including a load with plastic items, coffee mugs and stainless steel pots.

The STEP partners sought to highlight potential areas where test standards could be improved. However, this work is far from comprehensive, as it was limited in scope and was focused on the most critical areas identified. With this report we wish to contribute to the improvement of test standards, specifically by pointing out deficiencies, thereby helping standards to keep pace with the evolution of the market and technology. We also hope to foster the development of methodologies that produce results which are more relevant to consumers and for use to support legislation.

This report is based on test plans presented in Annex I. These test plans include conducting the EHTS but then also go beyond them, to try and apply some more real-life conditions to the testing. In so doing, this work has revealed some of the shortcomings of these standards. We will consequently emphasise shortfalls, reality gaps and omissions in the EHTS, bearing in mind the need for repeatability, comparability and affordability.

10. Spiliotopoulos et al (2017) Bringing the home in the lab: consumer relevant testing for household electrical products, ECEEE SUMMER STUDY PROCEEDINGS 2017
11. Link to EN 62087:2016: https://infostore.saiglobal.com/store/details.aspx?ProductID=1841650
12. https://www.cenelec.eu/dyn/www/f?p=104:110:87332793808301::::FSP_ORG_ID,FSP_PROJECT,FSP_LANG_ID:1257245,51905,25
13. https://webstore.iec.ch/publication/21803
14. https://www.cenelec.eu/dyn/www/f?p=104:110:87332793808301::::FSP_ORG_ID,FSP_PROJECT,FSP_LANG_ID:1257245,42520,25
15. http://ec.europa.eu/growth/tools-databases/mandates/index.cfm?fuseaction=search.detail&id=466
16. https://webstore.iec.ch/publication/23625

THE LACK OF REPRESENTATIVENESS OF TEST STANDARDS

What is it?

By representativeness, we mean the ability of test standards to reflect real life usage and conditions. While it is impossible to fully replicate the behaviour of all consumers, it should be possible to document the most common usage patterns and make sure test standards stay as close as possible to those, by covering the most common situations that the product would normally be in.

Why is that important?

- If unrepresentative, test standards will fail to give useful results and information to consumers, leading to potential wrong decisions and unpredictable energy bills (e.g., a dishwasher is only tested in eco-mode, but only 18% of real-life usage is in this mode);
- Unrepresentative test standards produce unrealistic information and make policy setting more difficult (e.g. a baseline appliance consumption calculated with a flawed standard, and to which regulatory action will be compared to, will provide an erroneous picture);
- If unrepresentative, test standards may fail to support the regulations in incentivizing manufacturers to improve the efficiency of certain features of the products; and
- If unrepresentative, test standards may fail to demonstrate the benefits and shortcomings of certain characteristics of products, creating an uneven playing field and possibly making EHTS more susceptible to circumvention.

Main findings

Dishwashers:

The test standard requires assessment based on only 1 programme out of 30 to 50 possibilities

Most dishwashers offer 5 – 7 different programmes, including eco, normal, intensive, short/quick, and glass. Most of today's dishwasher models also offer an 'auto' programme, which is supposed to adapt the cleaning cycle based on the degree of soiling detected. These same dishwashers also offer 2 – 5 extra functions such as 'short' or 'extra drying' that can be combined with a separate pre-wash and a self-cleaning function. In some cases, programmes can be combined with several extra functions, meaning users can have 30 or more combinations. Unfortunately, the energy label declaration refers to only one of these. This disparity in terms of user options and the declared energy performance can lead to a discrepancy between expected and real energy consumption.

To illustrate the different consumption patterns linked to the various programmes and extra functionalities, the table below provides an indicative range of some of the power consumption levels measured relative to the standard test programme (Eco standard programme) which is used for determining the energy class of a dishwasher.

Model	Eco standard programme (reference for reporting)	Eco with extra functionalities programme	Auto programme	Intensive programme
DW Model A.	0.94	1.20 (28% higher)	1.00 (6% higher)	1.40 (49% higher)
DW Model B.	0.89	1.33 (49% higher)	1.11 (25% higher)	1.35 (51% higher)
DW Model C.	0.83	1.07 (29% higher)	0.93 (12% higher)	1.44 (73% higher)
Overall	Reference	28 to 49% higher	6 to 25% higher	49 to 73% higher

Table 1: Measured energy consumption in kilowatt-hours of different programmes (IEC standard load, simple soiling).

Note: as extra functionalities are specific to each model, we report in the Eco+ extra functionalities column the highest power consumption found in our testing work. The purpose is to illustrate the difference with the Eco standard measurement, and not to compare models.

It clearly appears that there are significant differences in energy consumption linked to extra functionalities or specific programmes. If used differently from the way a dishwasher is measured for the energy label declaration, our testing has found that the appliance can consume up to 70% more energy than declared.

The tested programme is not the most used programme

A recent University of Bonn consumer survey (Hook, Schmitz, Stamminger et al., 2015) showed that the Eco programme is used for only 19% of all dishwashing cycles. Household users were found to select the normal / regular 45°C, 50°C or 55°C programmes 22% of the time. The normal /regular 60°C or 65°C dishwasher cycle was selected 17% of the time and the Short / Quick programmes 11% of the time. Taking all the normal programmes together, these accounted for nearly 40% of all dishwasher cycles.

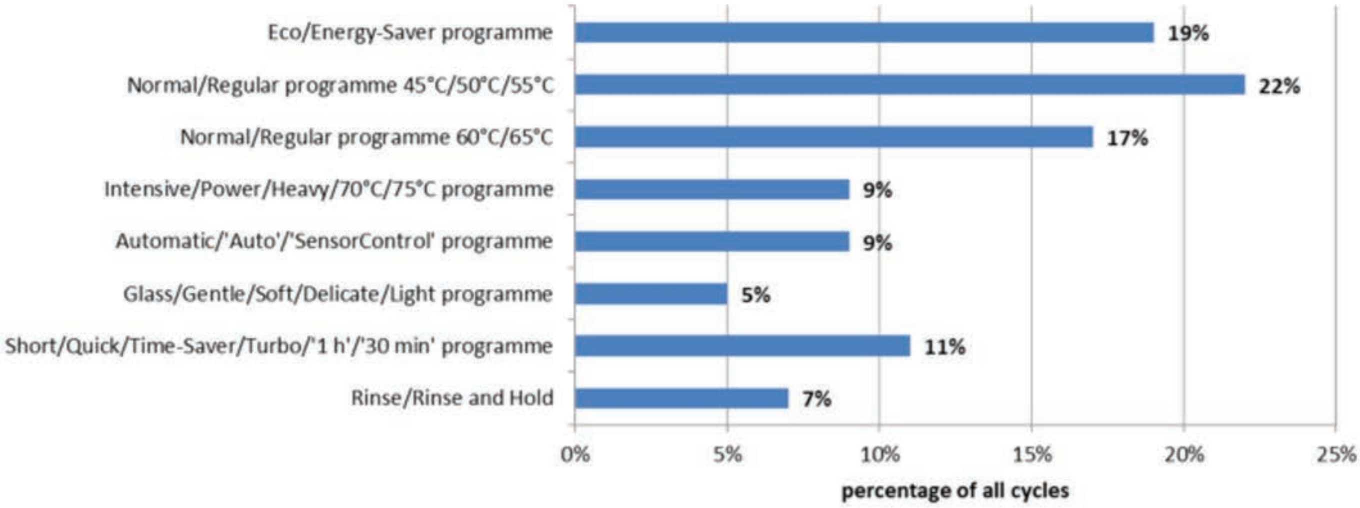


Figure 3: Dishwasher programmes used[17] Source: Hook, Schmitz, Stamminger et al., 2015

The STEP partners are concerned that the testing standard for dishwasher relies on a wash programme that is used less than 20% of the time by households across Europe. Our finding is that the energy label would be a much better consumer guide if it were based on the most commonly used 'normal' programme or combination of programmes.

Televisions

Outdated video test loop

Central to IEC 62087-2:2015 is a ten-minute standard video test loop for measuring the average power consumption of televisions. Unfortunately, this test clip does not represent what people watch on TV, and it has numerous cuts in the video content, not mirroring current broadcast patterns. Televisions have also become increasingly sophisticated over the last decade, adding features like ultra high definition (i.e. higher resolution), high dynamic range (i.e., more colours and greater contrast ratios), internet connectivity and complex software to manage and improve picture quality. However, the core element of test standard IEC 62087:2007 has not been modified – the ten minute video test loop used to measure average power consumption of the television – for 10 years.

The ten minute IEC 62087-2:2015 video test loop is characterised by over 100 separate video clips from around the world that are spliced together to form the loop. The resultant ten minute video was considered to be typical of average viewing globally, in terms of average picture level. Considering the fact that the previous international standard for measuring power consumption of a television involved a static picture power measurement, this video was a significant improvement. The IEC technical committee that prepared the video clip was not concerned with the number and frequency of the scene cuts. This resulted in the final video in IEC 62087:2007 being constructed with clips that last on average 2.3 seconds. In real live broadcast, video incorporating characteristics like this (including light and dark alternating clips) are rarely found, meaning that the IEC test video is somewhat unusual and is not representative of television broadcasting and videos. The problem with having video that has such distinctive characteristics is that it leaves the test standard vulnerable to misuse.

17. Dishwashing programmes used (percentage of all cleaning cycles) : 'Please indicate what kind of programme you use for washing your dishes and how often it is used.' Weighted, n = 3216 consumers.

For that reason, the STEP partners worked to develop a new video loop, designed to better reflect most common programmes broadcast on TV. STEP developed this new ten minute test video that measures energy consumption of HDR televisions and functions as an alternative test clip to the existing EN 62087:2016 test video clip.

This new ten-minute video sequence can serve as an alternative to the video sequence included with EN 62087:2016. The new video sequence offers the same average picture level as the IEC test sequence (34%), and has fewer scene cuts in the video, making it more representative of normal programme content (i.e., real world viewing material).

In Annex II of this report, the differences between the current standard test video clip and the new one developed by STEP are presented. In the table below we show the average measured power consumption for the same units tested with EN 62087:2016 (column 1) and the STEP video clip (column 2) loops in high definition (HD) resolution.

Television Tested	Measured average power using EN 62087:2016 HD	Measured average power using STEP HD	Percent change in average power
TV Model A.	105W	99W	-6%
TV Model B.	75W	75W	0%
TV Model C.	62W	68W	10%
TV Model D.	132W	134W	2%
TV Model E.	130W	191W	47%
TV Model F.	199W	206W	4%
TV Model G.	128W	129W	1%

Table 2. Measured average power consumption of seven televisions tested by STEP

The tests found that for one model the measured power is lower, for three models the measured power is similar, and for three models the measured power is more than 5 W higher. These results cannot be considered significant except for one model with a measured power difference of nearly 50%. We cannot extrapolate from this limited sample to the whole market, but the fact that such a significant difference was measured on one out of seven models shows that there is a need to adopt a more representative test loop and update the existing standard.

Fridge-freezers

The impact of door opening

The current testing standard for measuring the energy performance of domestic refrigerating appliances, EN 62552:2013, does not include any door opening during the test. The appliances are tested in a controlled environment with set room temperature and humidity conditions. Power consumption is measured with the appliance operating in a steady state with the door closed. This is also how energy consumption is measured in IEC 62552:2015. The standardisation community indicated that measuring power consumption in an ambient room temperature of 25°C, which is a higher temperature than usually found in kitchens, was a proxy for door opening. However, the STEP partners were concerned about this assumption and the reliability of a steady state condition to reflect real life conditions without including a door opening. Therefore, STEP designed a test programme that undertook to measure power consumption after both brief and long door openings.

A brief door opening was defined as opening the doors rapidly – just long enough for the appliance light to come on. If the appliance had two doors or more, these were opened simultaneously. For appliances that do not have a fresh food compartment the freezer door was opened for one second. .

A long door opening was defined as opening the doors multiple times over a 2 hour period intending to broadly mimic a relatively intensive episode in a normal home usage of the appliance[18]:

- Fresh food compartment door(s) opened for 10 seconds every 10 minutes.
- Freezer door(s) opened for 15 seconds every 30 minutes.

18. Based on Geppert, J (2011). Modelling of domestic refrigerators' energy consumption under real life conditions in Europe. Inaugural – Dissertation zur Erlangung des Grades, Hohen Landwirtschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität zu Bonn and Gemmell, A. (2017). Andrew Gemmell, Helen Foster, Busola Siyanbola, (BRE) and Judith Evans (RD&T). Study of Over-Consuming Household Cold Appliances – Field trial report, BRE for the Department of Business, Energy and Industrial Strategy, January 2017.

If an appliance had a fresh food compartment and a freezer, doors were opened simultaneously. If an appliance compartment had multiple doors they were opened simultaneously. The door opening duration was timed to last from the break of the door seal to being fully closed.

For the long door opening, the results show a visible impact on immediate power consumption (within two hours) for almost all of the models. For most, the variation was within what can be expected due to the physical door opening letting heat and humidity enter the appliances and the consequent energy needed to return the cooled space inside back to the set temperature. In a few cases, however, the long door opening test seems to have triggered more significant changes. The brief door opening had no noticeable impact on appliance temperatures and power consumption for eight of the ten appliances tested. However, for two of the tested units, even a short door opening was followed by what seemed to be interruptions of the stable state. In one case, the additional consumption was caused by an additional defrost episode happening just after the door opening (note: the test was then repeated with similar results). In the other case, a short door opening resulted in a significantly higher consumption and changes in the internal temperatures for over 24 hours.

For this single unit the average power consumption impact of door opening was significant both after the brief and long door opening, increasing by more than 30%. The magnitude of this increase is quite large, particularly having it sustained after a 12 hour period.

To better understand the impact of door openings in a consumer's home, we estimated the consumption of a 'normal day' usage pattern based on a usage pattern of two 12 hour loops of 2h of intermittent door openings followed by 10 hours of stable conditions with doors continuously closed. This scenario is closer to real life than the conditions currently defined in the standard test. We compared the average power consumed during this hypothetical day to the average power corresponding to the energy consumption measured following test standard EN 62552:2013, which does not prescribe door opening. The table below presents this indicative comparison. The last column shows how much lower the consumption, as measured following EN 62552:2013, is compared to the estimated consumption base on 12 hours loops with door openings.

Note: Such 12 hours loops were not performed during the test, but estimates were derived from the measurements made during the long door opening test[19].

19. Where door openings triggered changes of temperatures for more than 12 hours, average power of a 12 hour period with door openings was estimated based on the consumption during the 2 hours of door opening 2 hours following door opening, and during the time that the different temperature pattern had stabilised. This is to take into account the fact that in a succession of such 12 hours loops the appliances would never go back to the steady state that was the initial state in the test.

Model	Average power with door opening every 12 hours (Watts)	Average power from EN 62552:2013 measurements: % difference with door opening every 12 hours
F Model A.	27	-18%
F Model B.	22	-5%
F Model C.	38	-26%
F Model D.	22	<1%
RF Model E.	38	<1%
RF Model F.	37	-6%
RF Model G.	46	-8%
RF Model H.	53	-13%
RF Model I.	56	-3%
RF Model J.	77	-32%

Table 3. Estimated consumption with 2 hours of door openings every 12 hours, compared with measured consumption following EN62552:2013.

Note: F models refer to refrigerators alone, RF models refer to refrigerator-freezers.

With all the stated reservations associated with these estimates, the analysis nevertheless shows a significant difference in consumption that is above 10% for four appliances out of our sample of ten, whereas for some of the tested units there is no significant difference. This helps to further underpin the principle that product test standards should better reflect real life usage, and, for refrigerator-freezers, should take into account the effect of door opening in testing standards. It should be noted that the 2015 version of the IEC standard, supposed to replace the EN62552:2013, recommends testing the effect of door opening *"where the operation of a circumvention device is suspected"*. Clause 7 on circumvention devices indeed specifies that in such cases laboratories *"should subject the appliance to measures such as door openings or other appropriate actions in an attempt to detect presence and operation of any such devices"*. This is, however, only a recommendation and is only presented as relevant where circumvention is suspected.

While extrapolation to the whole European market cannot be done based on this sample of appliances, finding a case of significant deviation in energy consumption with door opening compared to the standardised test in such a small sample raises a potentially important issue. It is all the more significant that estimates presented above tend to confirm discrepancy in energy consumption between test standard measurement and 'normal day' usage patterns with door opening for four appliances out of ten, with significant differences between the models in the magnitude of the impacts of door openings. This is a very important finding for consumers and demonstrates that the test for the "determination of energy consumption" as described in the current EHTS and in IEC 62552:2015 is not sufficient to ensure a fair comparison of appliances. Results of the tests presented in this report lead us to recommend that the future European regulations and European harmonised standards should include measurements of the effect of one or more door openings for all domestic refrigeration appliances as part of the default standard tests related to energy performance, and not only in case of suspected circumvention.

> ***The results of this test on these three appliances show that the effect of loading on the performance (both energy efficiency and food preservation) differs widely between the refrigeration appliances tested.***

The impact of loading fridge-freezers

The testing standard EN 62552:2013 does not stipulate any realistic loading of interior space when testing the energy consumption of cold appliances[20]. The STEP project team decided to test a small sample of appliances to better understand the potential impact of loading. Three appliances were tested with 2 levels of loading in the fresh food compartment, holding all other conditions in EN 62552:2013 constant. Initially, pre-cooled (to 5±1°C) test packages were placed into the fresh food compartment of each appliance. This load was designed to be equivalent to 25% of the manufacturer's declared fresh food compartment storage volume. Once the temperatures and appliance operation were stable, a second test was conducted by placing 13.5 kg of test packs that had been equalised to 16±1°C. This load was based on data from Geppert (2011)[21] and was meant to represent a typical shopping trip. After the 13.5 kg load had stabilised in temperature and the appliance was back to a steady state, the total load (i.e., the 25% of declared fresh food compartment volume + 13.5 kg) was removed.

The results of this test on these three appliances show that the effect of loading on the performance (both energy efficiency and food preservation) differs widely between the refrigeration appliances tested. For one of the three models, the rise in temperature between before and after loading was as large as 9°C for one of the temperature sensors used for this test. For another model, the loading induced temperature variation inside the appliance was such that there were operation cycles of up to 8°C measured in some of the loading test packs.

In two out of the three models tested, the addition of pre-cooled packs and of an additional heat load caused temperatures to increase inside the appliances, although these increases were relatively minor in amplitude and/or limited in duration. For two of the three models, the difference of temperature between the different sensors in the fresh food compartment increased significantly. For one model, the energy consumption was slightly lower when the appliance was loaded and in one case it was significantly higher. In all three cases, within 6 to 12 hours of unloading, the temperatures and energy consumption went back to what it had been before the first load was placed in the appliance.

In all cases the operation of the refrigerator-freezer was significantly impacted by the loading, even after stabilisation. This could affect the level and distribution of temperatures reached in the different compartment, energy consumption, or both. Thus, the results of this test suggest that, like for door openings, a testing scheme that reflects the magnitude of the impact of loading would help ensure that the declared performance is not significantly different from what can be expected in real life.

20. In this version of the standard the freezer compartment is fully loaded but the fresh food compartment is empty.
21. https://www.researchgate.net/publication/263247009_Modelling_of_domestic_refrigerators%27_energy_consumption_under_real_life_conditions_in_Europe

THE LACK OF SUITABILITY AND TIME-RELEVANCE OF TEST STANDARDS

What is it?

By suitability we mean the ability of the test standard, and therefore consumer label, to take into account technical features of products affecting energy performance. While we can acknowledge that test standards cannot cover each and every possible specific functionality offered by each single product, the test standards should not neglect mainstream significant features that affect a whole product group.

Why is that important?

- If not comprehensive enough, test standards miss some sources of energy and resource consumption, or saving opportunities linked to appliance features (e.g. a mainstream functionality of appliances is not covered by the test standard);
- If not *dynamic/smart* enough, test standards neglect important factors influencing behaviours and energy consumption, thus undermining a proper assessment of the performances (e.g. a testing out of the box without first updating the appliance's software, which can influence performance).

Main findings

Dishwashers

One size does not fit all

There is no automatic correlation between energy class and energy consumption in programmes other than the standard programme, i.e. the Eco programme. Because dishwashers are only required to be tested in the Eco programme to declare the energy consumption and class, there is no evidence that the use of different programmes or additional functions ("short", "extra drying" etc.) different from the tested Eco programme are also reflected by the label class. For example, it is possible that an appliance in highest energy class may perform highly efficiently in all programmes, but it is also possible that there is no correlation between an energy class and the energy performance in other programmes than the tested one. This is an open question we cannot answer based on the tests performed in this project. What can be stated is that the standard test is ineffective in considering the variety of different wash programmes, and so it cannot be determined if a claimed efficiency refer to all programmes offered by the appliance or only by the eco-programme without extra functionalities alone.

Televisions

Mainstream functionalities not considered in the test standard

High dynamic range

A recent technological improvement in televisions is called high dynamic range (HDR). This system offers improvements in two areas: (1) HDR provides a greater contrast ratio (designed to take advantage of the full luminance range of a television) and (2) HDR provides a wider and richer range of colours. These two advantages mean that the displays offer more colour, much brighter whites and deeper, darker blacks, giving televisions a more 'dynamic' look, hence the name 'high dynamic range'.

HDR televisions are becoming mainstream as more and more media is being prepared in this format. According to IHS Markit, global demand for HDR televisions reached over 4 million units in 2016 and will grow to over 30 million by 2020[23], a growth rate of 65%. However, as the STEP partners noted EN 62087:2016, the test standard for measuring the power consumption of televisions, does not measure power when televisions use HDR format.

STEP created a new video test clip in three different resolutions: high definition (HD), ultra high definition (UHD)[24] and ultra high definition with HDR (UHD-HDR). We found the largest differences between the declared power and the measured power when the televisions were using the STEP UHD-HDR test video clip. Three of the seven models tested did not change significantly, but four of the models exhibited an increase in power consumption of between 32% and 130% when switching from playing HD content to UHD-HDR.

By not requiring the measurement of energy consumption in HDR mode, the current standard is ineffective at capturing the actual power consumption consumers experience in their homes when watching HDR content. The EHTS (EN 62087:2016) test video clip fails to trigger the HDR feature on an HDR TV, thus leading to lower power measurement. As HDR is a growing trend, the STEP partners felt that the standardisation community needs to make a revision to the test standard to add a procedure for measuring power consumption when playing HDR video content.

Automatic Brightness control (ABC)

ABC is an energy saving feature whereby a television uses a built-in light sensor to detect the ambient level of light in a room, and then adjusts the brightness of the screen to provide a more appropriate and comfortable viewing state. ABC is based on the principle that as light levels in a room are decreased, the screen does not need to be as bright for the same level of viewing acuity and perceived contrast ratio.

22. See: https://en.wikipedia.org/wiki/High-dynamic-range_video
23. Link: http://news.ihsmarkit.com/press-release/technology/hdr-tv-shipments-will-grow-43m-units-2016-hdr-compatible-market-more-four-t
24. Link to Wikipedia info on UHD: https://en.wikipedia.org/wiki/Ultra-high-definition_television

While overall the benefit to viewers, both in terms of comfort and energy savings, is well understood, there is no standardised test method for the reliable and repeatable measurement of a television's performance that incorporates ABC. For this reason, and to avoid any variance in power measurements, the current test standard requires that ABC is disabled prior to testing. However, STEP partners are concerned that by not testing or recognising this energy saving feature, the standard test neglects some energy saving potential, making the consumer label less reliable, and does not reward manufacturers for offering ABC. In our small sample of seven televisions, five were equipped with ABC, and we observed the power savings associated with this feature of 32 to 76% when varying the ambient room light levels from over 300 lux to less than 2 lux.

Models Tested*	Wattage Measured (>300 to <2 lux)	Power Savings
TV Model A	108W to 64W	41%
TV Model B	82W to 55W	33%
TV Model C	71W to 48W	32%
TV Model E	251W to 162W	35%
TV Model F	439W to 106W	76%

Table 4: potential power savings linked to ABC

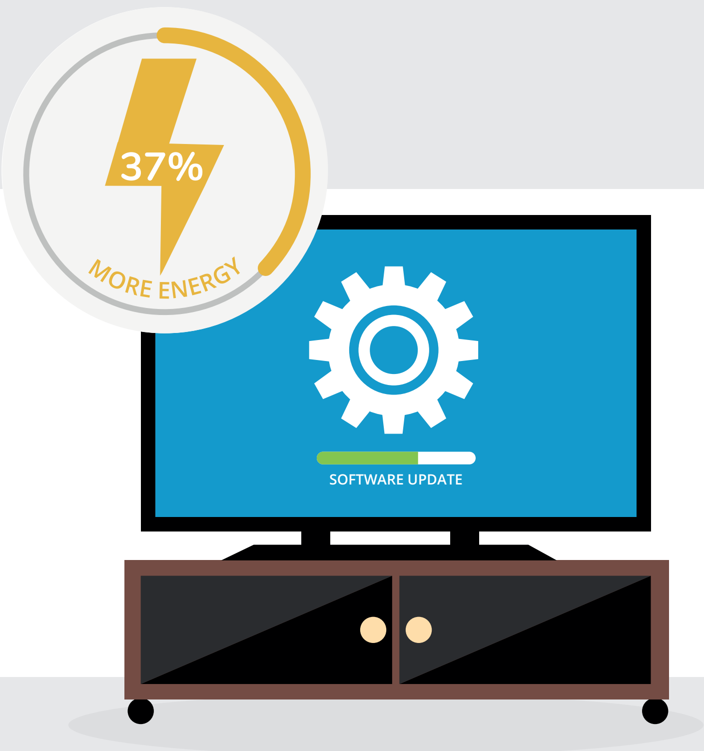
A new methodology to test ABC is suggested that could become a discussion baseline to update the standard test. An overview of this new ABC test methodology are presented in Annex III.

Difference between shipped settings and after software update

Smart, internet-connected televisions are now becoming a default feature. However, the standard test used today does not take this into account, meaning the television is tested 'out of the box' (i.e., as shipped). However, once connected to the Internet, the television will check for updated software and may ask the user for permission to download and install the new software. After the new software is installed, the television could have an entirely new default setting and could impact energy consumption.

This issue was observed for three of the seven television models tested when playing the same clip (UHD-HDR) before and after the software update. The graph below shows the change in power consumption over the ten minute STEP UHD-HDR video clip for one of the models. The increase in energy consumption was approximately 31% to 37%. Out of 7 models tested, the fact that three models had an increase in power consumption after a software update when measured with the STEP UHD-HDR test clip is significant.

The current test standard, and therefore the consumer label, does not capture the impact of software updates. One possible way for policymakers to protect consumers from the risk of increased power consumption following a software update is to incorporate into law that market surveillance authorities should download the very latest software updates prior to testing the product for compliance. This approach would ensure that if the manufacturer does develop software updates and made those available to consumers, the manufacturer would not be changing the power consumption of the product compared to the level at the time of purchase.



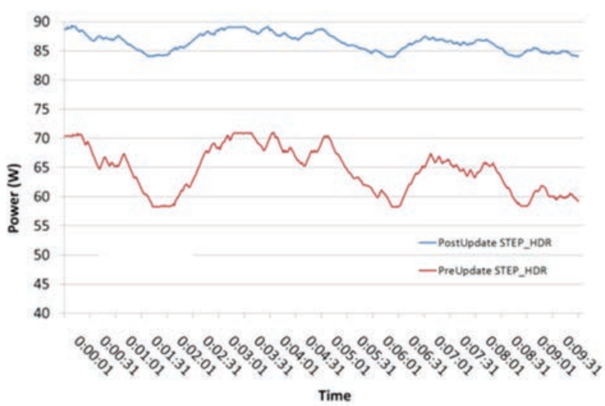


Figure 4: Change in HDR Power Consumption after Software Update (example for one model tested)

THE AMBIGUITY OF TEST STANDARDS

What is it?

Test standards should be clear in the methodology they demand of laboratory technicians. While we recognise that there will always be a place for interpretation on how to apply a standard test, this should be strictly limited and not affect the measurement results in any significant way.

Why is that important?

- If test standards cannot be applied to certain models within a product group, the comparison is flawed (e.g., a standard test cannot apply reliably with certain technologies,
- If test standards require measurement without enough precision, declared performances appear poorly verifiable (e.g., when measuring a parameter that can vary according to the shape of the appliance and no clear guidance is provided on how to handle that diversity)

Main findings

Fridge-freezers

We undertook the measurement of the storage volume of appliances, as this parameter is necessary to the calculation of the Energy Efficiency Index (EEI) on which label classes are based. The overall result using the current standard methodology was that only 4 out of ten appliances tested were within the authorized 3% tolerance margin when comparing measured values to declared values. In view of the influence of storage volume on energy label class ranking, this lack of guidance and precision in measuring storage volume renders the current testing standards too unpractical to grant the fair assessment of energy performance and fair comparisons between models.

Guidance on how to measure storage volume is not clear enough. Instructions in the current standard EN 6552:2013 do not describe precisely enough how to precisely measure storage volume. As a consequence, while the measurement of storage value is crucial in determining the energy class of a product, the methodology for storage volume measurement in this standard is quite complex and open to interpretation on level of measurement required.

For example, no guidance is provided on how detailed the measurements should be and whether small indentations and mouldings can be ignored. Additionally, it is almost impossible, without three dimensional mapping or a CAD diagram of the appliance, to accurately measure the volume of protrusions.

Figure 5: Examples of moulding that can be found in domestic refrigerating appliances. *Note: the refrigerator model shown in this figure is selected from the internet for illustrative purposes and does not necessary represent one of the models tested under the STEP project.*

It seems, however, that aligning the future European harmonised standard on an improved version of IEC 62552:2015 would help remediate the issue. The IEC 62552:2015 volume measurement procedure is clearer and therefore more appropriate for volume declaration under Ecodesign and energy labelling regulations than the storage volume measurement methodology currently described in EN 62552:2013. But issue remains, as the standards offer no guidance in case of a poor installation of moulding. For example, partitions or fittings can vary between apparently identical parts of the appliance. STEP observed in several appliances that mouldings could sometimes be moved or were not correctly located. These tended to generate difficulties for interpreting measurements.

Difficulty to apply the current standard with inverter compressor technology

Two out of the ten appliances tested use inverter compressor technology. This may become a mainstream feature in a near future, due to its higher efficiency potential. However, the official current standard appears unsuitable for models with this technology. These take a considerable amount of time to reach a stable energy consumption pattern, even when temperatures are stable. Therefore, these appliances can take considerable time to test, and the results can be less reproducible. The current standard procedure is not clear on how to address inverter technologies.

THE LACK OF ADEQUATE CONSUMER INFORMATION AND EMPOWERMENT

What is it?

By adequate information, we mean the provision of clear information on energy consumption and other parameters (water, noise etc) of certain settings/ programmes and features offered by the product. Of course, product instructions should not attempt to document comprehensively all possible behaviours. But useful information should not be neglected and wording should not be confusing.

Why is that important?

- If regulations do not set useful information requirements, users may miss certain behaviours can have an impact on energy consumption or other parameters (e.g.,: no information on the consequences of switching from default shipped settings or programme)
- If test standards do not prevent the use of confusing wording/messages, the users may misinterpret the impact of certain functionalities.

Main findings

Dishwashers

Confusing formulations

The eco programme, supposed to be the most efficient programme in terms of combined energy and water for normally soiled tableware, is used in official tests of dishwashers and is referred to as the 'standard' programme. However, there are also other programmes named 'normal', a more popular choice for consumers. They are understandably confused about which programme to use, as 'normal' clearly implies it is the programme to be used on a daily basis. This raises the clear question, why make the 'normal' option different from the tested one? This is especially true as tests show that cleaning performance is perfectly adequate under the eco/standard-programme. This may lead people to not use the eco programme by default although it seems this is the expected behaviour by making eco programme the one determining the label class.

Short programmes can be misleading. Some people think they are more efficient, a fact illustrated in the graph below. This belief is not always true.

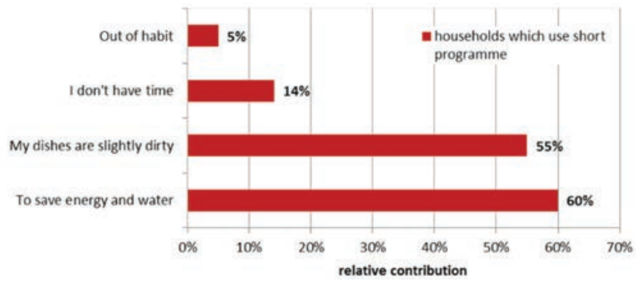


Figure 6: A survey of the reasons people give for selecting short programmes.[25] Source: Hook, Schmitz, Stamminger et al., 2015

Some short programmes do consume less than the eco/ standard programme. Out of the three models STEP tested, two have a short programme actually showing less power consumption than the Eco programme in our measurements. The table below illustrates the energy consumption of short programmes, without extra functionalities, compared to the Eco programme for two models.

Model	Eco standard programme without extra functionalities	Short programme without extra functionalities
DW Model A.	0.94	N/A
DW Model B.	0.89	0.72
DW Model C.	0.83	0.64

Table 5: Comparison of measured power consumption between Eco-standard programme and short programme without extra functionalities in kilowatt-hours (IEC defined load, simple soiling). Note that cleaning performance was not assessed and we cannot state if the dishes were similarly cleaned when using ecostandard and short programme.

Poor information on the impact of extra functionalities

As the Eco programme can be combined with several optional functions, tests were conducted on the Eco programme plus extra functions with the full load and simple soiling, in accordance with IEC 60436. If the Eco programme is combined with extra functions such as 'short' or 'extra drying', which are common features offered by dishwashers, we observed a 30% to 50% higher energy consumption. Manuals do inform users about the impact on energy consumption of some, but not all of these extra functions.

25. Consumers' motives for using the short programme ; multiple choices ('What is your reason for using short programmes ?'), weighted n = 1540.

There are no defined standards for delivering the appropriate information to users about the impact of extra functions. This exacerbates further the possible discrepancy between declared performance on the label (based on the Eco programme without extra functions) and a more typical real-life usage programme, which may involve other programmes and extra functions.

The STEP testing found that auto programmes consume more energy than standard Eco programmes. This can be confusing as a consumer may think the auto programme is more efficient than the eco/standard programme because it implies that it adapts to the load and soiling at the time of use. Thus it is not so clear when and why consumers should use the auto programme.

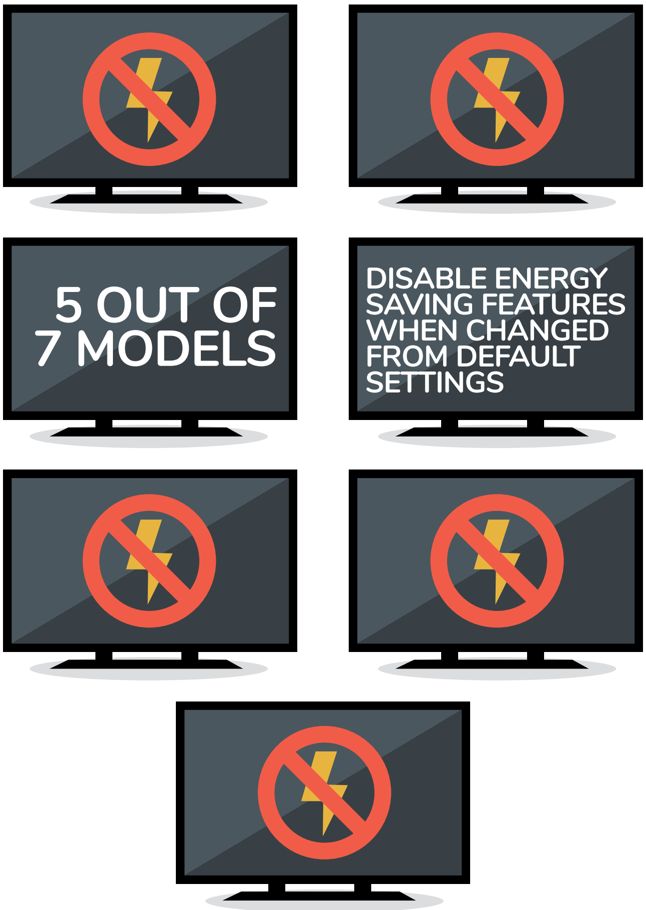
Model	Eco standard programme without extra functionalities	Auto programme without extra functionalities
DW Model A.	0.94	1.00
DW Model B.	0.89	1.11
DW Model C.	0.83	0.93

Table 6: Comparison of measured power consumption between Eco/standard programme and Auto programme without extra functionalities in kilowatt-hours per cycle (IEC defined load, simple soiling).

Televisions

No warning on consequences of changing picture settings

Televisions today offer users several different picture modes which manufacturers have adjusted to improve the viewing experience for certain content, such as sports, cinema, etc. Televisions also allow users to make their own custom adjustments to colour, contrast and other picture settings. The STEP partners sought to understand whether changes to the picture mode or picture settings would result in changes to the average power used by the television. The testing also sought to check whether the user was informed of the consequences of changing settings, such as whether energy saving features that were enabled would be disabled by modifying the picture settings.



The STEP testing found that 5 out of the 7 televisions tested disabled energy savings features when changing from the default picture setting. In two cases, certain energy saving features were deactivated and greyed-out, so it was not possible to re-enable those features without a factory reset. In all 5 cases, the user was not informed of the energy saving feature being disabled, nor the impact this selection has on energy consumption. Consequently, the user was not informed that changing settings may lead to a difference between the energy consumption of the television and the energy efficiency class declaration. Policymakers and the standardisation community need to look at establishing procedures and requirements to ensure that users are informed about the energy impact of picture settings they are making prior to these changes being made. In addition, users should always be able to undo setting changes if they are unhappy with the new setting and to return to their previous one.

Fridge-freezers

The effect of temperature/thermostat setting

Cold appliances currently on the market can be equipped with temperature settings using a numerical index such as 1, 2, 3, 4, 5, a button with arrows pointing to + or – signs, or numerical temperatures displayed on a small screen. STEP testing confirmed our expectation that the colder the setting, the more energy is used, however we are concerned that the consumer may not be well informed about this impact. With six empty fridge-freezers tested, STEP found that a one degree lower temperature can lead to an increase in energy consumption of 4 to 8%. This may be different in real life when appliances are normally loaded and doors occasionally opened. But such information about the consequences of setting changes are not properly and systematically communicated to consumers.

In addition, it is worth noting that the median temperature setting (e.g. 3, if a numerical scale from 1 to 5 is provided) corresponds to a target temperature that may differ from one model to another. The median setting does not always correspond to the same targeted temperature for freezing compartments or fresh food compartments, meaning the consumer is not empowered to set the appliance to the right temperature, or the desired one. It is therefore problematic when temperature controls fail to give a precise and reliable indication of the target temperature. This raises some concerns about the transparency and user-friendliness of temperature controls, and the impact thereof on energy and food. Controls that use a precise and reliable indication of the target temperature are also less problematic because they avoid any misinterpretation of the scale. Indeed, in some cases it could be that users understand such scales as referring to degrees Celsius rather than an inverted relative temperature scale.

Energy saving or special mode

Some of the tested models offered modes with names that seemed to imply that they should be enabled to save energy, for example when the user leaves for a long period of time. The STEP project decided to investigate the effect of these special modes on power consumption. With initial conditions set as for an energy consumption test under EN 62552:2013, none of the modes were found to deliver significant energy savings. One was even found to trigger significantly higher consumption than the steady state, with the average power consumption almost 50% higher in the 24h following the activation of the special mode than in the 24h before.

Although the sample in this test is small, these few examples show that it would not be appropriate to define rules for energy labelling that would benefit refrigerating appliances with special modes, even if the names of those modes seem to indicate that they will save energy. It would, on the other hand, seem relevant to provide clear information to the users about what special modes would change in the operation of the appliances, and when and why they should be enabled.

CASES DESERVING FURTHER INVESTIGATION TO CHECK PROPER COMPLIANCE AND RISK OF CIRCUMVENTION

As mentioned, the purpose of this report is not to check compliance with performance requirements and declared energy class of tested models, and the procedure we followed cannot justify any statement with regards to non-compliance, notably because only one and not three additional models were tested.

All televisions had average power consumptions that were roughly the same when tested with the EN 62087:2016 test video (which is HD format) and the STEP HD test video. There was, however, one model where the average power consumption of the television increased by 47% when playing the STEP HD video. The magnitude of this increase is concerning, given that the average picture level of the two HD test video loops is the same, and in HD mode, the power should be approximately the same (as it was for the other six models). Given the magnitude of this power increase, and the fact that none of the other models exhibited the same behaviour, this model stands out as potentially detecting and adjusting its behaviour to reduce average power consumption when measured with the EN 62087:2016 test video clip.

Out of the fridge-freezers tests, some discrepancies were also found between the values declared for energy class, and those measured. When comparing *declared* performance based on *declared* storage volume and *declared* power consumption with *measured* performance based on *measured* storage volume and *measured* power, five out of ten appliances were found to be at least one energy class less efficient. However, two appliances were in fact found to have a higher class than declared. While inaccurate, this is not so problematic for users. In addition, because of the inability of the testing standard to measure with utmost precision the storage volume, we maintain a conservative interpretation of the results and compare performance based on *declared* storage volume and *declared* power versus performance based on *declared* storage and measured power. This means accepting the declared value for storage volume due to the impracticalities mentioned above. Taking this conservative perspective, we still detected two fridge-freezers declaring an energy class higher than measured, even including a 10% tolerance margin applied to our measurement. Ten percent is the official tolerance margin applied to verification procedure by public authorities, but should not be used by manufacturers or importers when they declare product performance.[26]

26. For more information, see Commission Regulation (EU) 2016/2282 of 30 November 2016 amending Regulations (EC) No 1275/2008, (EC) No 107/2009, (EC) No 278/2009, (EC) No 640/2009, (EC) No 641/2009, (EC) No 642/2009, (EC) No 643/2009, (EU) No 1015/2010, (EU) No 1016/2010, (EU) No 327/2011, (EU) No 206/2012, (EU) No 547/2012, (EU) No 932/2012, (EU) No 617/2013, (EU) No 666/2013, (EU) No 813/2013, (EU) No 814/2013, (EU) No 66/2014, (EU) No 548/2014, (EU) No 1253/2014, (EU) 2015/1095, (EU) 2015/1185, (EU) 2015/1188, (EU) 2015/1189 and (EU) 2016/2281 with regard to the use of tolerances in verification procedures (Text with EEA relevance), published in the OJEU on 20 December 2016 and effective on 9 January 2017. Link : http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.346.01.0051.01.ENG

	Declared volume and consumption**		Declared volume, measured consumption**	
	Calculated label class	Declared label class	Calculated label class	With 10% tolerance
RF Model F.	A+	A+	A	A
RF Model I.	A	A+	A	A

Table 7: Difference between declared and measured performance for two fridge-freezer appliances

Note: For model RF Model I a different energy class was even noticed when calculating the energy class based on declared storage volume and declared power consumption.

It is possible that individual units can be defective or damaged during transport, or simply unrepresentative due to production variability. For these reasons, the authorities test additional units if the first one is found to not meet the criteria defined in the relevant European regulations. The STEP project did not test more than one unit of any model, thus the discrepancies described above were based on tests of a single unit. These two fridge-freezers may deserve further investigation by national market surveillance authorities and details have been passed on accordingly.

PART II:
RECOMMENDATIONS TO IMPROVE PRODUCT TESTING

RECOMMENDATIONS ON KEY IDENTIFIED ISSUES

More representative tests

1. Ensuring standard tests reflect real-life usage patterns and conditions

Having found that (1) dishwashers only test in Eco mode; (2) televisions use a test video that is different from normal broadcast content; and (3) refrigerators are tested without any door openings or internal loading; it appears that there is a need for standardised tests to better capture how people use products and appliances. We do not expect that all possible use cases should be tested, and tests should try to avoid very complicated procedures when measuring energy performance. However, we also find that testing appliances in overly artificial, unrealistic conditions leads to major discrepancies between declared energy consumption and what people are likely to experience in their homes. The system today seems to be incentivising manufacturers to optimise energy performance around those modes that are subject to the standardised test, without necessarily ensuring those optimisations are also experienced in the other modes and settings.

Consequently, repeatability and reproducibility are essential, but not the only criteria needed for the appropriateness of test methodologies in standards used when implementation the ecodesign and energy labelling regulations. Likewise, improvements in the methodologies are not necessarily associated with a significant additional testing cost, but even when that is that case, the cost parameter should be examined within the context of a wider picture. A small increase in testing cost can potentially deliver wider economic benefits for the consumer and the economy, and be necessary for Ecodesign to effectively ensure a level playing field. We therefore call for the wide and systematic consideration of representativeness, as a key criterion when formulating test methodologies for product standards. This has been identified as an important issue by decision makers, and is reflected in-legal provisions of the newly revised EU Energy Labelling Framework Directive. A reference is made to this approach in article 13.3: *"Harmonised standards shall aim to simulate real-life usage as far as possible while maintaining a standard test method."*

2.Combining testing of different main programmes

When purchasing an appliance or product, a consumer expects the energy rating to be reflective of how that model will perform relative to others in all of the different modes. Thus, policymakers could decide to combine several different programmes and functions into one measured energy metric for regulation. This would enable an assessment of different functions and aspects, which are combined in real life operation, and therefore should also be combined in testing as much as possible. This approach will also reduce the risk of test standard circumvention.

More suitability

1. Ensure test standards are keeping up with technological evolution of appliances

Consider timely process to update test standards with additional or combined testing, targeting technologies and functionalities identified as likely to go mainstream.

The case is very clear for televisions, where mainstream functionalities such as high dynamic range (potentially increasing power consumption) and automatic brightness control (potentially optimising power consumption) are totally neglected in today's standard. Our proposed improvement could be a generic recommendation for all product groups that are experiencing rapid technological evolution, for which standard tests run the risk of quick obsolescence.

Beyond fast evolution, the case of inverter compressor technology in fridge-freezers also illustrates the issue. The technology cannot be easily measured with existing testing protocol.

2. Software updates

Market surveillance authorities may wish to ensure they update product software to the very latest version before conducting compliance tests, as this is most likely to happen in the home when consumers are offered a software update from the manufacturer. If the appliance increases its power consumption after a software update, it may no longer be in compliance with its declaration of the energy label class. In extreme cases, it may fail Ecodesign regulations that would otherwise have prevented its sale.

Clearer guidance

1. Labs need clear instructions on key parameters

As illustrated in the case of storage volume of fridge-freezers, those criteria that influence energy performance and energy label class declaration should be free from ambiguities on how lab technicians conduct measurements of crucial parameters. Systematic guidance should be provided where such risks are detected. Such guidance should be produced as an integral part of the official test standard or be adopted using a fast track approach in case risks are identified once the standard is already in place. While this may already happen, it may be important to define a default approach if in doubt. For example, there could be transitional measurement methods or guidance released by the European Commission when certain issues are identified which require time to be integrated in the testing standard.

Better information

1. Energy labels should reflect 'normal' programmes

The testing standard should be based on the 'normal' or most used programme/s, but never those associated with Eco wording Eco should be reserved for energy saving programmes beyond what is reflected on the energy label.

This would not only lead to potential further savings, but also clarify matters for users. As revealed by tests on dishwashers, naming the standard wash cycle Eco can cause people to think it is not an appropriate programme for everyday use, when it is. Furthermore, the Eco programme was not always found to be the most energy saving programme.

2. Impact of settings turned on or off

The standardisation community may consider making it a requirement to provide more explanations to consumers on when activating or not certain feature and what will be the impact on power consumption. As seen in the case of special modes for fridge-freezers, such modes do not necessarily deliver savings compared to normal settings as used in the test standard even though their names imply that they save energy. Besides, they should be activated in clearly explained circumstances in order to avoid food waste. Conversely, it should become mandatory to warn the consumer before switching off any energy-savings / eco-features. This notification should inform the consumer of increased energy consumption and running cost, and require them to accept the change or not.

Of course, if no impact on energy performance can be linked to the appliance update, such a request for consumer information becomes less essential, though it could remain an option. It is worth noting that formulations adopted under the new EU Energy Labelling Framework Directive reflect this concern. Article 3.4 reads:

"Once a unit of a model is in service, suppliers shall request explicit consent from the customer regarding any changes intended to be introduced to the unit in service by means of updates that would be to the detriment of the parameters of the energy efficiency label for the unit, as defined by the relevant delegated act. The customer shall be informed of the objective of the update and of the changes in the parameters, including any change in the label class. For a period proportionate to the average lifespan of the product, the customer shall be given the option to refuse the update without avoidable loss of functionality".

It remains to be seen how this will be implemented product-by-product.

A NEW APPROACH TO TESTING AND ENFORCEMENT

Setting limits on variability

As seen with dishwashers, televisions and fridges, one of the challenges of setting relevant test standards is to make them reflective of real life usage and conditions. Consumers can choose from dozens of different combinations of programmes and settings for washing dishes, but only one programme is tested. We can use our televisions in many different viewing modes and formats, but only 'out of the box' settings are tested. We can operate our freezers at -18°C or -25°C without necessarily seeing the difference and without getting information with regards consequences on energy consumption; and we certainly open doors of our fridges on a daily basis, but nonetheless they are tested without any door opening. All these simplifications to the measurement methods are made to ensure good repeatability, shorter testing periods and lower costs for testing.

To address this challenge, we suggest an approach that has been developed in the automotive sector for the purposes of testing real life emissions. Under this proposal, automobiles would still be tested according to the EHTS (i.e., in a lab under very specific conditions), but in addition to that, they would also be tested on the road, with a wide range of driving patterns. The new real-driving emissions test standard for cars[27] defines the boundary conditions of what the test laboratory expert can do, giving the examiner flexibility to conduct their own driving pattern and measure the associated emissions.

We propose to apply a similar principle to appliances, such that manufacturers would establish a maximum deviation from the standard measured power consumption. The market surveillance authority would conduct the EHTS, and then their own unique test and measure the power consumption for both, determining whether the energy consumption of the unique test is contained within the declared maximum deviation. To avoid extreme outcomes, programmes or scenarios of little relevance to real-life would not be included in the range of possible unique tests.

For example, a dishwasher would be tested in normal mode and then again in an energy intensive mode with heavily soiled dishes. The product examiner would then check that the energy consumption in the heavily soiled cycle remained within the defined limits. For a television, it could mean that the TV is tested both 'out of the box' and then again in 'sports' or 'theatre' mode with a high dynamic range video clip and in a brightly lit room. Under these conditions, the non-standardised test must not consume more than a defined power limit compared to the standard protocol. For a

27. http://europa.eu/rapid/press-release_IP-15-5945_en.htm

refrigerator, it could mean that the appliance is tested both with the standard protocol and then again with the doors being opened and the storage space loaded. Again, under this non-standardised test scenario, the measured power would not exceed the declared consumption by more than a declared percentage.

It will of course be very important to define the range of allowable power consumption levels, and this will depend on each product category. When verified, a product would then be declared compliant if it not only respects the declared performance and energy class, but also remains within the set limits for maximum variation.

Finally, it has to be stressed that the proposal for the establishment of such maximum deviations and side tests does not imply that the main standardised methodologies should not be improved with regards to representativeness (see section 2-1-1 above). Otherwise, the risk is that we end up with products designed in a way which programmes and modes have very small deviations amongst each other, but with the main standardised test remaining unrealistic and irrelevant for consumers (i.e. methods that optimise deviations between programmes, but maintain low ambition).

Better consumption information, notably for Eco / efficient programmes

We suggest that manufacturers deliver more precise information on how much their different programmes and possible specific functions/settings consume, notably how much the most energy saving mode/setting/programme saves compared to the one(s) used to declare the energy class, and in which conditions this option could be most relevant. This low consuming option could be for example indicated as Eco and related information could be formulated as the Eco programme consumes X% less energy than displayed on the label or declared. This would ensure consumers are not confused with Eco or similar wording (*is it the programme to be used by default or only when one wants to save energy?*), and this would unleash further savings without disappointing consumers because when to use this Eco/most efficient programmes will be explained.

Towards a more effective enforcement system

In addition to the aforementioned approaches, the STEP partners call for a smarter and more effective system of market surveillance.

1. Thinking of enforcement as a European matter

Ecodesign and energy labelling are single European market instruments. It is logical and appropriate that there should be appropriate EU level support for market surveillance and enforcement. This would not necessarily mean transferring the competency for controlling the market from the member states to the Commission. However, more could be done to promote coordination and share testing costs, results and best experience among national authorities, making the best use of resources. *The Commission could act as coordinator, ensuring information flows between countries and eventually supporting testing of product groups that remain untested by national authorities. This principle is in line with the newly revised Energy Labelling Directive in Article 8.2: "The Commission shall encourage and support cooperation and the exchange of information on market surveillance of energy labelling regarding products covered by this Regulation between national authorities of the Member States that are responsible for market surveillance or in charge of the control of products entering the Union market and between them and the Commission, inter alia by involving more strongly the Administrative Co-operation Group (ADCO) on Ecodesign and Energy Labelling. Such exchanges of information shall also be conducted when test results indicate that the producer is in compliance with the relevant law". And Article 8.4 continues with "The Commission shall, in cooperation with the Administrative Co-operation Group (ADCO) on Ecodesign and Energy Labelling, elaborate guidelines for the enforcement of this Regulation, in particular as regards best practices of product testing and the sharing of information between national market surveillance authorities and the Commission."*

2. Improving rates of compliance

In view of the numerous models placed on the market and the limited resources dedicated to enforcement by market surveillance authorities, better EU coordination will help improve compliance, but it may not solve the problem by itself. STEP partners have identified the following options to improve compliance:

(a) Member state governments could adopt higher sanctions against non-compliant companies as well as naming and shaming unscrupulous companies that wilfully mislead consumers about product energy performance. In most countries across Europe, non-compliance is not widely communicated to the public. Instead, market surveillance authorities contact the concerned industry directly and negotiate a solution bilaterally. This approach limits the reputational damage for the non-compliant company, and whatever the sanctions, they are never as dissuasive as wide publication of non-compliant models and brands.

(b) Another approach could be to establish a system of consumer compensation when products are found to be non-compliant. Today, if consumers become aware of non-compliance, they can only expect compensation during the legal warranty period associated with the product, which is generally two years. But if as it was suggested by the European Parliament during the revision of the Energy Labelling Framework, it could be required systematic consumer compensation applicable throughout a product's lifetime. This would trigger a more systematic and wider publication of non compliance, and drastically increase the financial and reputational consequences for companies at fault. This would act as a powerful dissuasive instrument and make suppliers much more attentive to ensuring proper compliance.[28]

Conclusions

The STEP project aimed at identifying opportunities for capturing even greater CO2 emission reductions in Europe from product standards and labelling through (1) identifying issues and potential failures in standardised product testing; (2) documenting discrepancies between real-life and declared performance; and (3) suggesting improvements to the standardisation community to address any discrepancies or declaration problems which are found. Through the study of standardised tests for televisions, fridge-freezers and dishwashers, the STEP partners found four overarching problems relating to test standards:

- **A lack of representativeness of test standards** – more realistic test methods could result in more accurate quantification of performance, more accurate labelling and, potentially, more accurate energy savings potential assessment;
- **A lack of suitability of test standards** – certain characteristics of product performance are not measured effectively or simply do not have a test method;
- **An ambiguity of test standards** – existing standards are not precise enough or less adequate given how product technologies have changed; and
- **A lack of user information** – product modes can vary and the impact on energy consumption is not properly or clearly communicated.

To address these concerns, the STEP partners identified potential improvements for these different aspects that we hope will be adopted by the European standardisation community and policymakers when revising the relevant regulations and standards. A central tenet behind our recommendations is that societal stakeholder organisations, such as those representing the environment, consumers or workers, bring invaluable expertise and defend interests which are otherwise under-represented (if at all) in the standardisation system. A balanced representation of interests and effective participation of societal stakeholders involved in standardisation will contribute to increasing the quality of the standards and the legitimacy of the process. Therefore, a balanced representation of stakeholders with a diversity of interests should be sought, and their effective participation should be ensured.

In addition to achieving such balance, the European Commission should exercise more control over the specifications to be followed in test methodologies. Policy makers in charge of EES&L can have more control over the definition of the test standards than the European Commission does. For example Australia put in place a process to speed-up and control the delivery of some test methods. In the United States, the regulatory authority for energy-efficiency, the Department of Energy, has control over the development of both test methods and performance regulations, and develops them both through an open and transparent public process. [29]

Overall, it is clear that the standardisation process needs to anticipate technological evolution of appliances and the way we use them, using a timely approach to adjust and revise standards. In line with the recently revised energy labelling legislation, we recommend requiring test standards are better reflect real life conditions. Furthermore, we call for a better and systematic investigation and documentation of user behaviour in the regulatory review and preparatory studies, and for that information to be used as early as possible by the standardisation community.

The STEP partners outline a new approach to testing products based on the principle of being transparent about the deviations of energy use related to multiple modes and settings of the product from the measured performance under a standardised setting and the declaration. This idea offers a more systematic approach to conveying information on expected savings linked to using most efficient/eco modes compared to what's declared on the label. We think this approach could be relevant for several product groups covered under ecodesign and energy labelling.

28. It is worth noting that such legal provision was rejected by the member states during the discussion on Energy Label reform, but the idea is still worth exploring through reforming the DIRECTIVE 1999/44/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees.
29. A 2014 CLASP report presents more details on these examples and outlines suggested improvements and a few longer term scenarios of how the standardisation process could evolve. Developing Measurement Methods for EU Ecodesign and Energy Labelling Measures, a discussion paper. Edouard Toulouse, published by CLASP Europe. February 2014. http://clasp.ngo/Resources/Resources/PublicationLibrary/2014/Alignment-of-EU-Test-Procedures-and-SL-Regulations.

ANNEXES

ANNEX I: TESTING STEPS AND PROCEDURES FOLLOWED FOR EACH PRODUCT GROUP

Televisions

The STEP testing of televisions focused on six types of tests to assess their impact on energy consumption:

1) Average power consumption measurements
2) Software updates and supplementary power measurements
3) Networked standby power consumption measurements
4) Television menus and energy saving features
5) Automatic brightness control (ABC) power measurements
6) Sound level and power consumption measurements

The procedure we followed is summarised in the table below:

Test Plan Item	Description of Task
1. Average power consumption measurements	Without connecting to the Internet, power up the television, accept all the default settings and measure the average power consumption using the ten minute test video loops following a standard 'out of the box' set-up. Document which Eco-features are enabled by default. Disable Automatic Brightness Control (ABC) if enabled; check to ensure other Eco-features were not impacted by that step. Conduct average power measurement tests with the following four test loops: - EN 62087:2016 (standard) – HD format - New STEP Test Loop – HD format - New STEP Test Loop – UHD format - New STEP Test Loop – UHD HDR format In addition, conduct measurement of the 'out of the box' luminance levels, in as-delivered settings.
2. Software updates and supplementary power measurements	Connect the television to the Internet, and direct it to download any available software updates. Record the new software version and document which Eco-feature settings have changed. Conduct the four power tests as listed in item #1 in this table.
3. Networked standby power consumption measurements	After verifying the television is connected to the Internet, put it into standby mode and log the power consumption for 18 hours, in one second intervals.
4. Television menus and energy saving features	Study the default viewing mode and how it changes in response to menu changes. Make an adjustment to any of the non-Eco-feature picture settings. Document which Eco-features remain enabled and which have been switched off. If any are switched off, document whether the user was informed of them being disabled.
5. Automatic brightness control (ABC) power measurements	With the updated software (as it would be in the home), conduct the ABC test runs, to assess the power consumption changes and television display luminance response characteristics in ambient light levels measured at the ABC sensor from 300 lux down to 12 lux. Analyse the graphically presented ABC control characteristics for both power and luminance
6. Sound level and power consumption measurements	Ascertain the energy-impact of the TV's integrated sound system, using the testing methodology for Audio systems as set out in EN 62087-6:2016 Evaluate the proportion of on-mode power consumed by the sound system of low-end and high-end product samples at a standardised listening level

Fridge-Freezers

The STEP testing of fridge-freezers focused on 6 types of tests to assess their impact on energy consumption:

1. Measurement of interior volume using EN 62552:2013 and IEC 62552:2015
2. Comparison with energy and label declarations (10 appliances), using EN 62552:2013
3. Door openings
4. Temperature changes
5. Influence of load
6. Energy saving modes

The procedure followed is recapitulated in the table below.

Test Plan Item	Description of Task
1. Measurement of interior volume using: EN 62552:2013 IEC 62552:2015	The interior volume of 10 appliances is measured according to EN 62552:2013 *Household refrigerating appliances – Characteristics and test methods section 7-Determination of linear dimensions, volumes and areas, and IEC 62552-3, Edition 1.0 2015-02 Household refrigerating appliances – Characteristics and test methods – Part 3: Energy consumption and volume section 4 Applicable test steps for determination of energy and volume -4.8 Volume determination*
2. Comparison with energy and label declarations (10 appliances), EN 62552:2013	Appliances are tested following EN 62552:2013 section 15 Energy consumption test, 8 General test conditions and 13 Testing storage temperatures. The energy use measured through this test will be compared to the energy use declared by the appliance manufacturer and will serve as a baseline for the interpretation of the results of the rest of the energy consumption tests.
3. Door openings	This test includes both a short door opening test and a long door opening test. The short door opening shows what controls may be activated by the opening itself, keeping the temperature and humidity changes to a minimum. The long door opening test shows the impact on energy consumption of an opening regime that resembles episodes of normal use.
4. Temperature changes	Additional appliance thermostat settings – the four freezers and four of the fridge-freezers are be tested at several additional thermostat settings to analyse the impact of those different settings on energy consumption and the link between the changes in the settings and the resulting changes of temperature in each compartment.

Continued...

Test Plan Item	Description of Task
5. Influence of load	All other conditions of EN 62552:2013 being respected, the fresh food compartments of three of the refrigerator-freezers is loaded with a simulated food product to determine the impact of stable load and of additional heat load on temperature and energy performance. Initially a pre-cooled (to 5±1°C) set of tylose test packs were loaded in the fresh food compartment of each appliance. This load was designed to be equivalent to 25% of the manufacturers declared fresh food compartment volume. The load of tylose test packs was distributed on and between the shelves and crisper sections as evenly as possible. The sensors fitted for the EN 62552:2013 remained in place and no additional sensors were placed in the loaded packs. Once temperatures and appliance operation were stable, a second test was carried out. This consisted of loading 13.5 kg of tylose test packs that had been equalised to 16±1°C, in addition to the first set of packs. This load was based on data from Geppert (2011) and was meant to represent a typical shopping trip. It was designed to be equivalent to adding 550 kJ, assuming that the appliances had to reduce the temperature of the tylose packs from 16 to 5°C. The packs were distributed as evenly as possible within the fresh food compartment but were not placed in crisper sections or doors.
6. Energy saving modes	Some of the purchased refrigerators have energy saving modes. The energy consumption of those appliances is tested following as these modes are activated. For that we apply the process as described in EN 62552:2013 although of course not respecting the target temperatures.

Dishwashers

The STEP testing of dishwashers focused on seven types of tests to assess their impact on energy consumption:

1. different programmes
2. combination with extra functions
3. different dish load
4. half load
5. different soiling
6. water with different hardness
7. with hot fill

The procedure followed is recapitulated in the table below

Test Plan Item	Description of Task
1. Different programmes	For assessing the performance of dishwashers according to the Energy Label and Ecodesign, only the Eco programme is tested. A 2015 consumer survey however revealed that for more than 80% of all cycles other programmes are used. Therefore, all programmes that had shown to be relevant for users were tested. For each of the three models that were selected, 'user-relevant' programmes were selected to for the test. Those are Eco, Intensive, Auto, Short and additional specific programmes depending on the model.
2. Combination with extra functions	The three dishwasher models offer extra functions that can be combined with all or most programmes. Some of these functions can also be combined with each other. In the instructions booklet manufacturers state that extra functions can change duration (and consumption) of the programmes. All extra functions available on the three dishwasher models were tested in combination with different programmes. Not all extra functions can be combined with any programme, some programmes also cannot be combined with any extra function. The possible combinations cannot be found in the manuals, but had to be defined once the test models had arrived at the lab. **Combination with the Eco programme was of highest interest, because it might change the consumption values of the Label programme.**

Continued...

Test Plan Item	Description of Task
3. Different dish load	The standard load in the current measurement standard (EN50242:2008) contains specific plates, cups, glasses, spoons, forks, etc. For each number of place settings, the number of each of these items is defined. Theoretically it is thinkable that a dishwasher can detect the weight of the standard load and thus recognize a test situation. Therefore, the load type and weight was varied. The new version of the European standard, aligned with the IEC standard, is introducing new items to the test load (mugs, pots, bowls, plastic items). The new load is supposed to better reflect typical load at peoples' homes. The new standard dishes were available at the lab, so the future standard dishes were used to vary the load type. Manufacturers often provide detail advice to test institutes on how the dishes are to be loaded into the appliance. Since these instructions were not available for the new standard load, testing engineers loaded the dishwashers under test according to the information that was available. There are larger items in the new test load, but the overall weight is slightly lower than that of the current standard load. Since with the new loading scheme, a slightly lower total mass of dishes needs to be heated up, this could lead to a slightly lower energy consumption. If instead the energy consumption with the new load showed to be clearly higher than with the current standard load, this might indicate an adaptation to the standard load. Since the new load is considered to be more consumer-relevant, more cycles were run with the new load (39) than the current standard load (22).
4. Half load	All three dishwasher models were tested with full load, but also with half load. One cycle was run with no load at all. Expected results were that the lower the mass of the dishes, the lower the energy consumption would be. The half load tests were aiming at two questions: ▪ Might some dishwashers be detecting the standard load and adapting their behaviour to achieve better results when being under test? If the energy consumption with half load were not lower than with full standard load, this could indicate detection and adaption happening. ▪ How much is the energy consumption reduced if dishwashers are only half filled? Expectations are that the reduction is less than 50%, and not filling the dishwasher fully is not an efficient was of use.

Continued...

Test Plan Item	Description of Task
5. Different soiling	According to the standard, dishes are heavily soiled, and the soil is dried. Apart from a few tests with standard soiling, 'simple' soiling was used for most test runs. For the simple soiling, a mixture of 15g of spinach, 15g of minced meat and 7.4g of margarine was applied to the inside of the dishwasher's door. In addition to this, one of the glasses contained 150ml of frozen milk. A glass of frozen milk is also used for acoustic test according to IEC 60436:2015. The objective of adding this to the simple soiling was to have turbidity sensors, if present, activated for some time during a cycle, because the milk would take some time to melt. The mixture of spinach, minced meat and margarine on the other hand would activate any sensor already in the pre-wash phase, while being washed away quite quickly. Heating up and melting the frozen milk uses some extra energy, approximately 0.015 kWh (55 kJ[30]). This simple soiling was used for most of the test runs, because the time-consuming soiling is an important cost factor. Using a simpler soiling for most of the tests allowed for more test runs being performed, and thus investigating the impact on dishwashing energy consumption of more parameters.
6. Water with different hardness	The preparation of water used for testing appliances is prescribed in EN 60734:2012 Household electrical appliances – Performance – Water for testing. For dishwashers, the water quality is defined regarding total hardness, the ratio of Calcium/Magnesium ions, and the maximum concentration of chloride, as well as iron, copper and manganese ions. For one test run per model, the water hardness was varied and softer water instead of standard test water was used. The dishwasher models' water hardness setting was not changed for this. First, this setting normally only influences the frequency regeneration occurs (more often with harder water), but not energy use. Second, changing this setting may become active after a number of cycles only. So, for just one cycle it would not make sense. This test was investigating if some dishwasher models might be detecting a test situation by recognising the standard water. If the energy consumption with soft water was higher that expected based on other test results, this could indicate a possible detection and adaptation.
7. with hot fill	Two of the dishwasher models, models B and C, can be connected to a hot water tank. If a household's water is heated with renewables (e.g. solar power), hot fill can save electricity. For the two models suitable for hot fill, hot fill was tested with the Eco and one other programme. For the model C, additionally the 'SolarSave' programme was tested in combination with hot fill. For all hot fill tests, 45°C hot water was used. These tests were aiming at quantifying the energy saving potential of hot fill.

30. The specific energy needed to melt ice is 333 kJ/kg. Melting of 150ml of milk thus requires roughly 50 kJ. The specific heat capacity of ice is 2.06 kJ/(kg*K). To heat up the 150ml of milk from -18°C to 0°C thus needs an additional 5.5 kJ.

ANNEX 2: COMPARISON BETWEEN THE CURRENT STANDARD TEST VIDEO LOOP AND THE NEW VIDEO LOOP DEVELOPED BY STEP

STEP contracted an independent European test laboratory to obtain their expertise in developing a new video test loop that incorporates the same average picture level (APL) as the IEC standard. The laboratory engaged the services of professional video editor Gerrard Giorgi-Coll and broadcast / film cameraman Andrew Rix to film and produce a ten minute video, with the same video footage but offered in different resolutions / formats:

1) Ultra High Definition (UHD) / 4K video, High Dynamic Range (HDR)
2) Ultra High Definition (UHD) / 4K video (normal, without HDR)
3) High definition (HD) video
4) Standard definition (SD) video

This video, in all formats, maintains the same 34% APL, which was determined by IEC as the world average in 2007. The work involved the preparation of five video sequences of two minutes each, covering a range of different TV programing in lighting / image content and reflecting average TV/display product programme content. The RGB colour mapping formats are consistent with those used in broadcast and distribution media standards including the new Industry agreed (2016) standard for HDR TV.

In preparing the actual video clips, actors were employed to avoid copyright issues, and filming venues were hired such as a news/current affairs type studio, a drama set studio and a sports venue. Exterior locations including street scenes were also used for filming. A professional 4K HDR camera was used and special post production computer editing equipment hire was required. The resultant video contains scenes that are of a longer duration than those found in EN 62087:2016 and which are more typical of real world viewing material. The video was filmed in ultra high definition (UHD) and rendered down to produce a high definition (HD) version. Metadata was added to the UHD sequence to produce both an 8-bit and 10-bit HDR pattern.

Aspect	EN 62087:2016	STEP (new sequence)
Duration of video clip	10 minutes	10 minutes
Average Picture Level	34% APL	34% APL
Formats available	Standard Definition High Definition	Standard Definition High Definition Ultra High Definition (4K) UHD High Dynamic Range
Number of cuts in ten minute video clip	261 cuts	97 cuts
Average scene length between cuts	2.29 seconds	6.19 seconds

Table 8. Comparison of IEC 62087:2007 and the new STEP test video sequence

The table below shows some of the content of the video clip itself, which can be downloaded for free (without copyright) from the CLASP website.

Time (mm:ss)	Content description	Example Frames
00:00 to 02:00	Outdoor: Street and river scenes around London – outdoor (bright mid-bright lighting)	
02:00 to 03:00	Sports: Snooker hall – indoor (dark and mid-bright lighting)	
03:00 to 04:00	Sports: Tennis match indoor (bright lighting)	
04:00 to 05:00	Advertisement: fruit blender, indoor (mid-bright and bright lighting)	
05:00 to 06:00	Advertisement: car polish, outdoor (bright lighting)	
06:00 to 08:00	Drama: indoor bar/pub scene, (dark lighting)	
08:00 to 10:00	News: breakfast TV indoor studio interview scene (bright)	

Table 9. Descriptive summary of the STEP test video loop

Please note that although this test video loop was modelled after the EN 62087:2016 video clip in terms of length and average picture level, it is only intended to function as an interim test video loop to provide policy-makers with another data point in terms of average power consumption (i.e., alternative to EN 62087:2016) for displays. A new IEC test loop in the appropriate resolutions (i.e., SD, HD, UHD) and formats (i.e., non-HDR, HDR) is still urgently needed and should be developed through the normal standardisation channels.

ANNEX 3: SUGGESTION FOR A NEW METHODOLOGY TO TEST AUTOMATIC BRIGHTNESS CONTROL (ABC)

STEP worked with an independent European test laboratory to develop a new test methodology for measuring a television's power response with ABC enabled. This new ABC test methodology has four important advantages:

- Easily controllable levels of illuminance shone onto the television's ABC light sensor via projector;
- Provides excellent granularity in terms of light levels as the projector's light output is cycled from full brightness to black and then back up again;
- Methodology is very efficient from lab technician time perspective; and
- Results are highly accurate and repeatable.

The test set-up is depicted in Figure A, with labels to indicate the key pieces of equipment and components involved in the test setup. The television under test is called the "test sample", and two luminance meters are used – one that is mounted perpendicular to and at the centre of the display, and a second one next to the ABC sensor itself. The projector shines its light output directly onto the test sample and the light output is measured.

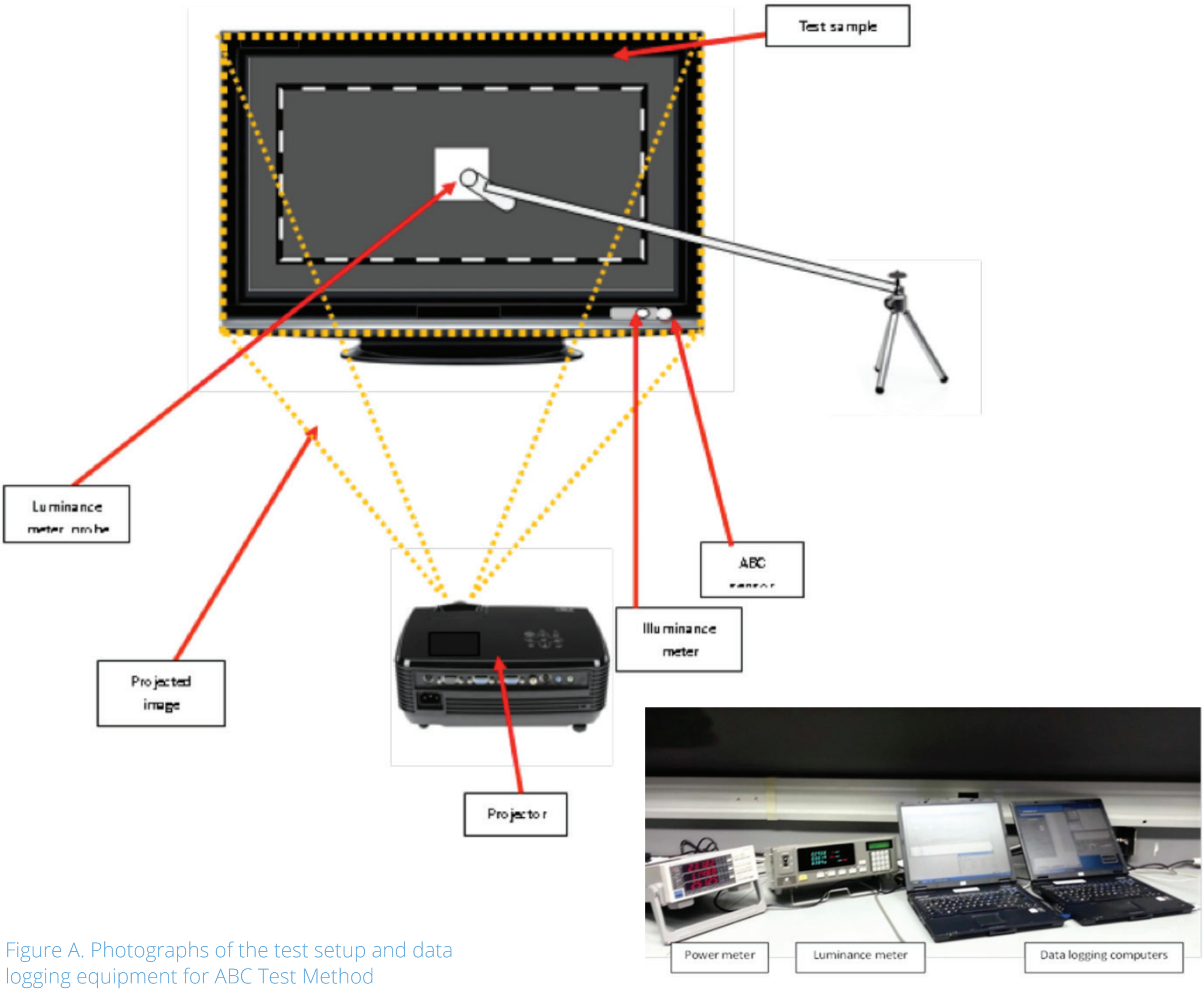


Figure A. Photographs of the test setup and data logging equipment for ABC Test Method

A summary of the ABC methodology is given below:

1) **Test Setup** – a television, projector and light sensors are setup as shown in the figure above, in a dark room and connected to the metering equipment. The television model being tested is connected to its standard AC input voltage via data logging power meter. A data logging contact colour analyser for display luminance measurements is mounted in the centre of the screen, within the boundaries of the European Broadcasting Union (EBU) test pattern peak white box. The positioning of the projector, which is perpendicular to the television model being tested, is not critical (unlike the current DOE lamp methodology); and the only precise alignment required is one to ensure that the illuminance meter is registering an illuminance value at the fixed test position very close to the value immediately in front of the ABC detector within 2 lux at approx. 100 lux illuminance) This is a very straightforward process on all the current test samples with ABC.

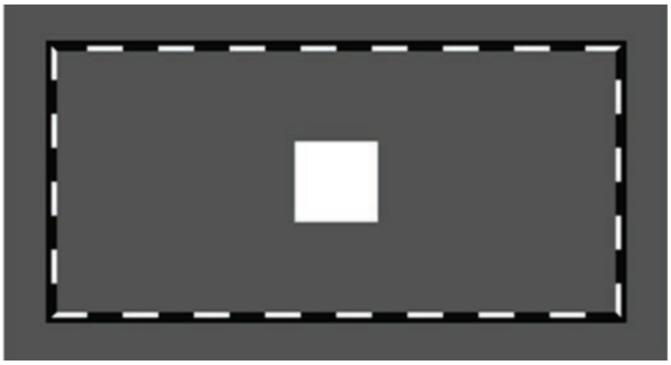


Figure B. EBU test pattern displayed continuously on television during test

2) **Adjust illuminance to >300 lux** – the distance between the projector and the UUT is then adjusted to provide an illuminance reading of more than 300 lux on a data logging chroma meter that is mounted next to (but not blocking) the light sensor of the UUT. Any reflective surfaces from trim or plastic that are part of the UUT are not masked, but all other surfaces between the projector and the UUT are dark and non-reflective.

3) **Vary illuminance and record data** – the projector is fed a slide show which consists of 39 slides that vary the light output of the projector from white (255, 255, 255) to black (0, 0, 0). These slides are played in an automated presentation mode with a five second duration for each slide in forward and then in reverse order. These slides are projected at the UUT in a dark room (i.e., no ambient light), and this has the effect of adjusting the illuminance measured by the television's light sensor. The black slide normally provides a lower illuminance at the light sensor of less than 2 lux. During the slide show presentation, the illuminance at the ABC detector and the power of the television are simultaneously logged in Excel with no technician involvement in the process. After reaching the end of the presentation – which ends on a black slide – the presentation is then run in reverse, back to the full white >300 lux white slide.