Independent Advisory Panel on characterising flavours in tobacco products

(IAP)

Methodology to support the decision whether a tobacco product has a characterising flavour

Application to cigarettes, roll-your-own tobacco and heated tobacco products

IAP approved this report on the 31/08/2023

This report provides the methodology to support the decision whether a tobacco product has a characterising flavour. The methodology is applicable to cigarettes, roll-your-own tobacco and heated tobacco products, in line with Directive 2014/40/EU, as amended by Commission Delegated Directive (EU) 2022/2100.

This methodology is prepared by the Independent Advisory Panel on characterising flavours in tobacco products, established by Commission Implementing Decision (EU) 2016/786 of 18 May 2016 laying down the procedure for the establishment and operation of an Independent Advisory Panel assisting Member States and the Commission in determining whether tobacco products have a characterising flavour, taking into consideration the input from the Technical Group and the best available practices in this field.

In line with Art. 19(1)(f) of the Commission Implementing Decision (EU) 2016/786 of 18 May 2016, the Commission will publish the methodology on the following dedicated website and provide a link to this website from the register of expert groups. For more information:

IAP website: https://ec.europa.eu/health/tobacco/products/characterising_flavours/panel_en

Register of Commission Expert Groups: https://ec.europa.eu/transparency/expert-groups-register/screen/expert-groups/consult?lang=en&groupID=3452

Keywords: tobacco products; boxed cigarette, roll-your-own tobacco, heated tobacco product, tobacco consumers; tobacco policy; characterising flavour; sensory analysis; chemical analysis

DISCLAIMER

The reports/opinions of the Independent Advisory Panel on characterising flavours in tobacco products ('IAP' or 'the panel') present the views of the independent scientists who are members of IAP. The opinions are published by the European Commission in the original language only.

ACKNOWLEDGMENTS

The substantial contribution of the Technical Group of Sensory and Chemical Assessors (EUREST-FLAVOUR consortium), which have proposed and developed the methodology and procedure for the product assessment is gratefully acknowledged.

About IAP

In line with the Tobacco Products Directive (2014/40/EU) and Commission Implementing Decision (EU) 2016/786, the Independent Advisory Panel on characterising flavours in tobacco products (IAP) has been established assisting Member States and the Commission in determining whether or not a tobacco product has a characterising flavour.

The IAP is tasked with issuing opinions on whether a tobacco product has a characterising flavour. The panel should also further specify and update, as appropriate, the methodology for the technical assessment of products as set out in Commission Implementing Decision (EU) 2016/786. In addition, the Commission may consult IAP on other matters relating to the assessment of characterising flavours. As appropriate, the panel shall request input from the technical group of sensory and chemical assessors that was established by the Commission via a public procurement procedure.

Scientific Committee members

Dr. Alberto Del Rio, Dr. Garmt Dijksterhuis, Dr. Jan van Amsterdam, Dr. Wouter Visser, Emmanuel Vanzeveren, Prof. Efthimios Zervas (Chair)

Contact:
European Commission
Directorate-General for Health and Food Safety
Directorate B - Public Health, Cancer and Health Security
Unit B4 - Health Promotion and Disease Prevention
Luxembourg (SANTE-IAP-CONSULT@ec.europa.eu)

About the Technical Group

The core members of the EUREST-FLAVOURS Consortium that worked on this report include scientists from the University of Crete, Greece [Dr. Constantine Vardavas (Coordinator), Christina Kyriakos, Ioanna Lagou, Valia Marou, Dr. Alexander Vardavas, Dr. Manolis Tzatzarakis, Prof. Aristidis Tsatsakis] and CARA Technology Limited, UK [Dr. Bill Simpson].

Catalogue number: EW-03-23-446-EN-N
ISBN: 978-92-68-09039-8
doi: 10.2875/837616

TABLE OF CONTENTS

1 TERMS, DEFINITIONS AND ABBREVIATIONS

1.1 Abbreviations

ANOVA: Analysis of variance

CATA: Check all that apply

CFTA: Check five that apply

CI: Confidence interval

EU: European Union

EU-CEG: EU Common Entry Gate

HS-SPME-GCMS: Headspace solid-phase microextraction gas chromatography-mass spectrometry

HTP: Heated tobacco product

IAP: Independent advisory panel[1]

LOD: Limit of detection (chemical analysis)

LOQ: Limit of quantification (chemical analysis)

LSD: Least significant difference

MS: Member States

PCA: Principal component analysis

RYO: Roll-your-own tobacco

[1] https://ec.europa.eu/health/tobacco/products/characterising_flavours/panel_en

1.2 General terms and definitions

Analysis of variance (ANOVA): A parametric statistical technique used to investigate the sources of variation in a data set. Typically used in sensory testing to investigate variation due to samples, assessors and other experimental variables[2].

Additive: A substance, other than tobacco, that is added to a tobacco product, a unit packet or to any outside packaging[3].

Anchor: A point on a sensory answering scale and/or a physical reference against which comparative judgements are made.

Aroma: The sensation produced when volatile compounds stimulate olfactory receptors in the nasal cavity[2] often reserved for retronasal olfactory perception.

Attribute: A single perceptible impression/sensory characteristic of a product/stimulus[2].

Assessor: The individual, sensory panellist, respondent, subject, and so on, giving a response[2].

Characterising flavour: A clearly noticeable smell or taste other than one of tobacco, resulting from an additive or a combination of additives, including, but not limited to, fruit, spice, herbs, alcohol, candy, menthol or vanilla, which is noticeable before or during the consumption of the tobacco product[3].

Chemical standards: Chemical compounds of known identity and purity.

Cigarette: A roll of tobacco, within a paper cylindrical container, that can be consumed via a combustion process[4].

Clearly noticeable: The term 'clearly noticeable' used in the TPD has no formally recognised meaning in sensory science and is a colloquial term. Here, a flavour is considered as 'clearly noticeable' if it satisfies one or both of the following two criteria: (1) It is distinctive from and contrasting to that of the odour characteristics of the tobacco in which it is present and (2) It is more intense than that which would be regarded as merely "noticeable". The determination of whether a test sample possesses a "clearly noticeable" flavour is based on statistically defined cut-off limits further described in this report as part of the sensory assessment rank rating technique.

Composite flavour: A composite flavour comprises the combined effect of one or more different attributes in such a way that these would be perceived by a consumer. For the purpose of determining the intensity of this flavour characteristic, individual attributes are considered jointly, as one composite flavour, rather than individual flavour elements. For example, a product may possess a composite flavour of "mint chocolate" comprising individual flavour attributes of spearmint, peppermint, burnt sugar, vanilla and dark chocolate. As a result, the number of potential composite flavours possible is extremely large.

Descriptive profiling: The chosen sensory assessment for expert sensory panel members to identify and rate the perceived intensity of individual odour attributes.

Flavour: The total of sensations resulting from stimulation of the senses in the oral and nasal cavities, namely taste, olfactory, and trigeminal receptors[2].

Flavouring: The additive that imparts smell and/or taste[3].

Heated Tobacco Product: A heated tobacco product is a novel tobacco product that is heated to produce an emission containing nicotine and other chemicals, which is then inhaled by user(s) and

[2] Kemp, SE, Hollowood, T and Hort, J (2009) Sensory Evaluation. A Practical Handbook. Wiley-Blackwell

[3] EU Tobacco Products Directive (TPD) 2014/40/EU; http://data.europa.eu/eli/dir/2014/40/oj

that, depending on its characteristics, is a smokeless tobacco product or a tobacco product for smoking[4].

Independent Advisory Panel (or 'IAP'): A panel composed of highly qualified, specialised and independent experts with relevant expertise in the fields of sensory, statistical and chemical analysis that Member States (MS) and the Commission may consult to provide opinions as to whether test products have a characterising flavour[5].

Nonparametric tests: Statistical tests that do not make assumptions about the underlying distribution of the population or nature of the scales used to collect the data[2].

Odour: For the purpose of sensory testing in this report, the assessment is confined to the 'orthonasal odour' characteristics of the product prior to smoking. Therefore, the smell or odour of the product is evaluated rather than the retronasal odour, taste or mouthfeel of the products. Therefore, references to the sensory testing speaks about the 'odour' of the products. The odour of tobacco, both prior to and during burning, comprises numerous individual odour attributes. Each of these individual attributes arises from the impact of specific chemical compounds on the human olfactory system.

Odour attribute: a label (*e.g.* 'honey', 'potato skin', 'prune') of a flavour compound, occurs at the level of the individual chemical compound (and isomer).

Olfaction (or smell): The sense of smell[2]. 'Olfaction' refers to both orthonasal olfaction (up the nose) and retronasal olfaction (down the nose, via the mouth).

Parametric tests: Statistical tests that assume that the data from the underlying populations is normally distributed or follows another known distribution[2].

Principal Component Analysis (PCA) space: For the use in the current context, the objective of this method is to summarize the information contained in the data set into a graphical representation of tobacco products and attributes by breaking it down into principal components that explain as much variance as possible.

Rank-rating: A technique in which all samples, diluted with reference products, in a set are first ranked in order of perceived intensity of a chosen attribute or composite flavour and then rated for perceived intensity of that attribute or composite flavour[2].

Reference products: Three sets of tobacco products, either boxed cigarettes, RYO tobacco, or respectively, with the following characteristics:

- they represent, to the extent possible, products on the EU-market that are considered not to have a characterising flavour,
- they have undergone sensory and chemical testing,
- and are used as a basis for the comparison with a test product undergoing assessment for whether it has a characterising flavour.

Roll-your-own (RYO) tobacco: Tobacco which can be used for making cigarettes by consumers or retail outlet[4].

Sensory panel leader: A trained sensory professional who is able to train a panel of assessors to generate valid, consistent and reliable data[2].

Sensory panel: A group of assessors selected and trained to make objective sensory judgements of test samples[2].

Sensory analysis/testing: Sensory, including smelling, analysis is an established scientific discipline that applies principles of experimental design and statistical analysis to assess and describe

[4] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022L2100

[5] Commission Implementing Decision (EU) 2016/786; http://data.europa.eu/eli/dec_impl/2016/786/oj

perceptions by the human senses, including smell, for the purpose of evaluating consumer products. It has been found to be a suitable method for producing valid, robust, reliable and reproducible results when assessing whether a tobacco product has a characterising flavour [2,6,7].

Technical group of sensory and chemical assessors ('the technical group' or TG): Set up by public procurement to assist the IAP with an assessment of the sensory and, where appropriate, chemical properties of the test product as part of the procedure laid down in Article 10 of Commission Implementing Decision (EU) 2016/786[5].

Test product: A product referred to the panel by a Member State or the Commission for an opinion as to whether or not it has a characterising flavour within the meaning of Article 7(1) of Directive 2014/40/EU[3].

Test product assessment report: For each product that IAP has requested the Technical Group to perform an assessment on, a report would be provided including complete sensory analysis data and results of chemical analysis data, as necessary.

Tobacco: Leaves and other natural processed or unprocessed parts of tobacco plants, including expanded and reconstituted tobacco[3].

Tobacco products: Products that can be consumed and consist, even partly, of tobacco, whether genetically modified or not[3].

1.3 Specific binary classification terms and definitions

Positive sample: is a sample that truly has a characterising flavour other than that of tobacco.

Negative sample: is a sample that truly does NOT have a characterising flavour other than that of tobacco.

True positive: is when the sample is positive and the test result is also positive.

True negative: is when the sample is negative and the test result is also negative.

False negative: is when the sample is positive, but the test result is negative.

False positive: is when the sample is negative, but the test result is positive.

[6] Lawless, HT and Heymann, H (2010) Sensory Evaluation of Food. Principles and Practices. Springer, 2nd edition, 596 pp.

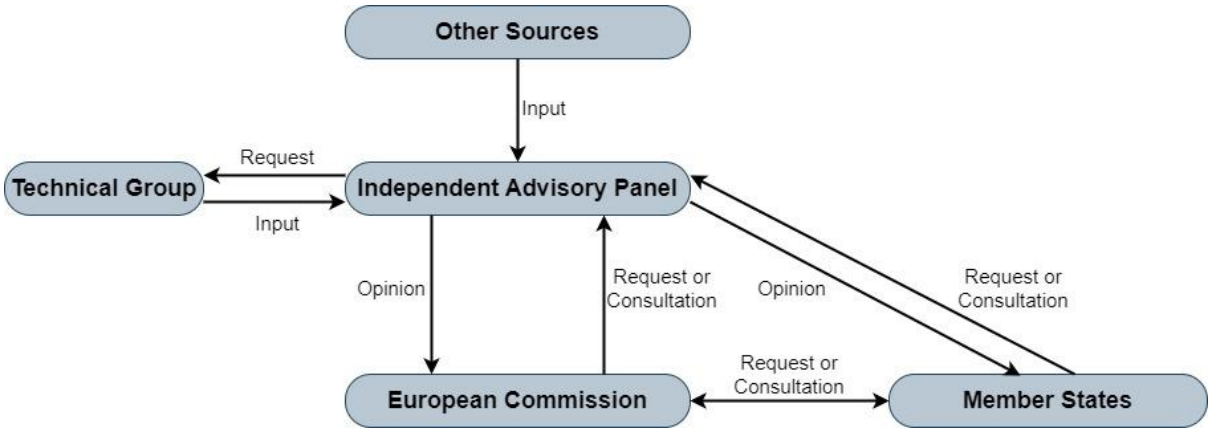
[7] Lawless, HT (2013) Quantitative Sensory Analysis. Psychophysics, Models and Intelligent Design. Wiley Blackwell, 404 pp.

2 BACKGROUND AND CONTEXT

Directive 2014/40/EU[3] of the European Parliament and of the Council of 3 April 2014 on the approximation of the laws, regulations and administrative provisions of European Union (EU) Member States (MS) concerning the manufacture, presentation and sale of tobacco and related products, prohibits in Article 7 the placing on the market of tobacco products (cigarettes and roll your own) with characterising flavours[3]. A characterising flavour in the Directive is defined as a "*clearly noticeable smell or taste other than one of tobacco, resulting from an additive or a combination of additives, including, but not limited to, fruit, spice, herbs, alcohol, candy, menthol or vanilla, which is noticeable before or during the consumption of the tobacco product*".[1]

Based on Directive 2014/40/EU Article 7, section (3) and (4), the Commission has developed two legal acts: Commission Implementing Decision (EU) 2016/786[5] of 18 May 2016, laying down the procedure for the establishment and operation of an independent advisory panel assisting the Member States and the Commission in determining whether tobacco products have a characterising flavour (Figure 1); and Commission Implementing Regulation (EU) 2016/779[8] of 18 May 2016, laying down uniform rules as regards the procedures for determining whether a tobacco product has a characterising flavour.

Figure 1. Overview of the procedure of the Commission Implementing Decision (EU) 2016/786



Article 7(4) of Directive 2014/40/EU provides that when determining whether a tobacco product has a characterising flavour, EU MS and the Commission may consult an independent advisory panel ('IAP' or 'the panel'). The procedures and operation of the IAP are outlined in Commission Implementing Decision (EU) 2016/786.

Commission Implementing Decision (EU) 2016/786 stipulates the IAP should be assisted by a technical group of sensory and chemical assessors ('TG' or 'the technical group') for the technical assessment of test products. The TG was established through a public procurement procedure, as part of Single Framework Contract (Chafea/2016/Health/36) to provide services to support the assessment of characterising flavours in tobacco products, commissioned to an external contractor (the "EUREST-FLAVOURS" Consortium). Commission Implementing Regulation (EU) 2016/786 specifies that the technical group should carry out sensory and chemical assessments based on a comparison of the smelling properties of the test product with those of reference products and where it is considered applicable, the sensory analysis can be complemented by chemical analysis of the product.

Sensory, including smelling, analysis is an established scientific discipline that applies principles of experimental design and statistical analysis to assess and describe perceptions of the human senses, including smell, for the purposes of evaluating consumer products. The analysis should be conducted

on the basis of an established methodology and that the panel shall specify and, as appropriate update the methodology for the technical assessment of test products. In developing the methodology, the panel shall take into consideration, as appropriate input from the technical group.

The Commission may request IAP for an opinion on whether a tobacco product has a characterising flavour. Where IAP considers it necessary for the purposes of providing an opinion, it shall request input from the Technical Group (TG) established in accordance with Article 12. In forming its opinion, the panel shall have regard to the information and data obtained from the technical group. It may also have regard to any other information at its disposal that it considers authoritative and relevant, including information resulting from reporting obligations pursuant to Article 5 of Directive 2014/40/EU[4].

Article 2(28) of the Tobacco Products Directive 2014/40/EU[8], defines the '*substantial change of circumstances*' as an increase of the sales volumes by product category by at least 10 % in at least five Member States (MS), based on sales data transmitted in accordance with Article 5(6) of the Directive or an increase of the level of prevalence of use in the under 25 years of age consumer group, by at least 5% in at least five MS for the respective product category, based on the Special Eurobarometer 385 report of May 2012 or equivalent prevalence studies.

In June 2022, the Commission published a Report[9] indicating that the sales volume of Heated Tobacco Products (HTPs) had increased by more than 10% in 10 MS and that HTPs accounted for 3.33% of the total sales volume. This finding triggered a substantial change of circumstances for HTPs. On this basis, the existing prohibition of selling tobacco products with a characterising flavour the ban was extended to HTPs. This extension was implemented through a delegated Directive[10] under Article 7(12) and Article 11(6) of Directive 2014/40/EU. Previously, this prohibition only applied to cigarettes and roll-your-own tobacco. Additionally, the Commission eliminated the option for Member States to exempt these products from the labelling requirements outlined in Article 9(2) and Article 10 of Directive 2014/40/EU.

This report provides the specifications of the established methodology used to provide an opinion to the Commission on whether a tobacco product (cigarette, roll-your-own, HTP) has a characterising flavour, based on input from the technical group.

2.1 Basis of the sensory assessment

Human perception of flavour is the combination of sensations resulting from stimulation of the chemical senses in the oral and nasal cavities, namely olfactory (*i.e.* smell), taste and trigeminal (i.e. mouthfeel) receptors. The approach for sensory analysis in the current context solely assesses the orthonasal odour characteristics (*i.e.* smelling up the nose) of the product prior to smoking. The chosen methodology for sensory analysis is designed to lead to results that are accurate, repeatable, and reproducible internally between sensory panel members, and also consistently repeatable between different sensory expert panels, and therefore are not specific to a particular sensory panel. The selected sensory testing procedures and practices, including aspects of the test facility, test

[8] Directive 2014/40/EU of the European Parliament and of the Council of 3 April 2014 on the approximation of the laws, regulations and administrative provisions of the Member States concerning the manufacture, presentation and sale of tobacco and related products and repealing Directive 2001/37/EC (OJ L 127, 29.4.2014, p. 1)

[9] Report from the Commission on the establishment of a substantial change of circumstances for heated tobacco products in line with Directive 2014/40/EU (COM/2022/279 final) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022DC0279

[10] Commission Delegated Directive (EU) 2022/2100 of 29 June 2022 amending Directive 2014/40/EU of the European Parliament and of the Council as regards the withdrawal of certain exemptions in respect of heated tobacco products (OJ L 283, 3.11.2022, p. 4–6)

design, test selection, test execution and analysis of results are based on well-established methodologies in the field of sensory evaluation, to be further described within this report [6,7,11,12,13,14].

The HETOC report[15] describes an approach in which flavour attributes are generated by the panellists and therefore sensory panel specific. Subsequently, the results of each assessment are sensory panel specific, and hence not straightforwardly comparable across sensory panels, with each sensory expert panel naming attributes and rating their intensity in a way which reflects the specifics of the training and the composition of its members.

The chosen methodology described in this report can be compared across sensory panels, *i.e.* the perception of all flavour attributes by sensory assessors and by consumers is driven by the presence of specific concentrations of individual flavour compounds. Furthermore, the sense of olfaction in humans is extremely specific, and a particular feature of olfaction is that the response of assessors to individual stereoisomers of the same compound differs both qualitatively (in terms of odour characteristic) and quantitatively (in terms of odour intensity). The human olfactory system is able to combine stimulation by several odourants to form a composite odour. Thus, they may use terms such as *fruity* or *floral* to describe mixtures of individual flavour compounds which impart individual fruity odours reminiscent of apples, blackberries or raisins, or individual floral odours, such as those resembling violets or roses.

Nevertheless, perception, recognition and identification of flavour compounds, mainly occurs at the level of the individual chemical compounds and mixtures thereof. The methodology further uses anchoring the results of the sensory evaluation of tobacco products to verifiable chemical standards. The attribute intensity results generated are therefore reflective of both the sample evaluated and the sensory panel that made the evaluation. Hence, for the application of Directive 2014/40/EU and to fit regulatory needs it is essential to anchor the results of the sensory evaluation of tobacco products to verifiable chemical standards. Chemical reference standards of known concentration and purity, and therefore known flavour character and intensity, can provide such anchoring. Therefore, reference standards have been used during the training of the sensory panel, to define the existence and intensity scales used by individual assessors and the sensory panel as a unity. This allows for a sensory panel to be trained using the same approach, using the same standards and vocabulary and hence these are applicable by both the industry and regulators, within other settings.

The chosen sensory assessment for sensory expert panel members to identify characterising flavour attributes, descriptive profiling, is a well-established method used in the field of sensory analysis. Attributes which are absent from individual samples are scored 0 by individual assessors. When a given attribute is present at the lowest possible intensity it is possible for the assessor to detect, it is assigned a score of 1. Intensity scores higher than one are assigned on a proportional basis (*i.e.* a score of four being twice the intensity of a score of two) with a maximum allowable score of ten. Results from the flavour profiling of a product are reproducible when sensory panellists are trained appropriately.

2.2 Basis of the chemical assessment

As described in the HETOC report, "Sensory testing is required to determine the limit of detection by humans beings. In addition, although objective chemical analysis identifies the individual compounds

[11] Chambers, E and Wolf, MB (1996) Sensory Testing Methods. American Society of Testing and Materials, 115 pp.

[12] Stone, H and Sidel, JL (1985) Sensory Evaluation Practices. Academic Press, 2nd edition, 338 pp.

[13] Kemp, SE, Hollowood, T and Hort, J (2009) Sensory Evaluation. A Practical Handbook. Wiley-Blackwell, 196 pp.

[14] Kilcast, D. (2010) Sensory Analysis for Food and Beverage Quality Control. A Practical Guide. Woodhead Publishing, 373 pp.

[15] HETOC report. RfS Chafea/2014/health/19. "Mapping of best practices and development of testing methods and procedures for identification of characterising flavours in tobacco products".

of tobacco, perhaps components at low concentrations combine to form a flavour that could only be detected and identified subjectively. Therefore, chemical analysis often is validated with sensory testing to determine flavours."

As described in the HETOC report, "Several analytical methods, such as gas chromatography (GC), infrared spectroscopy (IR), nuclear magnetic resonance (NMR), and mass spectrometry (MS), can be used to analyse tobacco ingredients. Since tobacco additives are often present at trace levels, analytical methods need to be sensitive, such as gas chromatography-mass spectrometer (GC-MS) methods"[13].

Furthermore, it is important to note that the chemical analyses can be performed with several methods: headspace (HS), solid-phase micro-extraction (SPME), solvent assisted flavour evaporation (SAFE), direct solvent extraction (DSE) and steam distillation (SDE) analysis.

Previous research has indicated that, through the application of specific chemical procedures it is possible to create a flavour profile for a tobacco product through the identification and possible quantification of chemical components in tobacco products. The chemical analyses so complement the product sensory analysis.

Overall, the HETOC report concluded that the approach for product testing should be based on a comparison of the smelling properties of the test products with those of reference products, complemented, as appropriate, by a chemical assessment of the product composition through chemical analyses. Their conclusion was to combine the use of an expert sensory panel with a chemical analysis assessment. The chemical approach is based on well-established methodologies for qualitative and quantitative analysis of tobacco additives in tobacco products[16,17,18,19].

[16] Krüsemann EJ, Visser WF, Cremers JW, et al. Identification of flavour additives in tobacco products to develop a flavour library. Tobacco Control 2018;27:105-111.

[17] Merckel C, Pragst F, Ratzinger A, Aebi B, Bernhard W, Sporkert F. Application of headspace solid phase microextraction to qualitative and quantitative analysis of tobacco additives in cigarettes. Journal of Chromatography A, 1116 (2006) 10–19

[18] Jiu Ai, Kenneth M. Taylor, Joseph G. Lisko, Hang Tran, Clifford H. Watson, Matthew R. Holman. Menthol Content in U.S. Marketed Cigarettes. Nicotine Tob Res. 2016 July ; 18(7): 1575–1580. doi:10.1093/ntr/ntv162.

[19] Lee C, Lee Y, Lee, JG, Buglass AJ. Development of a reduced pressure headspace solid-phase microextraction-gas chromatography/mass spectrometric (rpHSSPME-GC/MS) method and application to aroma analysis. Analytical Methods 7(16), 6504-6513, 2015.

3 ESTABLISHING THE SENSORY PANEL

3.1 Selection and recruitment of the sensory panel

Commission Implementing Decision (EU) 2016/786 refers to the composition of the technical group, which shall be composed of experts including, "*Sensory assessors recruited on the basis of their olfactory discrimination ability and their capacity to perceive, analyse and interpret smells, and who have reached the age of majority as laid down in applicable national legislation*". A systematic approach was taken to identify, recruit, evaluate and select the desired type of individuals and evaluate their potential to work well as a team.

Recruitment methodology comprised of:

(i) An advertising and outreach programme targeted at local individuals with strong roots in the local area
(ii) Provision of a briefing document and questionnaire to prospective candidates
(iii) Review of the information supplied by prospective candidates and selection of suitable candidates
(iv) Invitation of selected candidates to participate in a sensory aptitude assessment
(v) Initial evaluation of sensory aptitude assessments (three hours each)
(vi) Review of the aptitude assessment results and selection of the best performing candidates
(vii) Invitation of the best performing candidates to complete a personality questionnaire and attend a behavioural test session
(viii) Behavioural assessments (two hours) together with a briefing for candidates
(ix) Confirmation by candidate assessors of their willingness to undergo training and accept the roles if offered
(x) Final review of data, including advice from a behavioural psychologist concerning the outputs of the personality tests
(xi) Selection of candidate assessors for training
(xii) Approval of the candidate assessors by the Commission

The target profile of candidate assessors was based on a criterion of 11 mandatory characteristics and six desirable characteristics. Candidates who completed the application form and were not excluded from the process on the basis of the criteria outlined above were invited to attend a sensory aptitude assessment session. The primary goal of the sensory aptitude assessment phase of this task was to identify candidates who were best able to learn to recognize aroma compounds and use specific names to identify them when instructed by a skilled facilitator.

Success in such aptitude assessments depends on candidate assessors having:

- ✓ An excellent ability to follow both written and verbal instruction;
- ✓ An excellent sense of smell, with no significant impairments;
- ✓ An excellent ability to memorize the characteristics of each aroma in a short time;
- ✓ An excellent ability to discriminate one aroma from another, with a low probability of confusing similar aromas with one another;
- ✓ An excellent ability to resist potential distractions during the blind-coded tests.

In other approaches to assessor aptitude assessment, the aroma recognition skills of candidate assessors are tested without any prior training. The success of candidates in such tests thus depends exclusively on prior experience, rather than on them possessing the characteristics listed above.

Points achieved by individual candidate assessors in each of the three test sessions were combined and expressed as an Overall Score (%), with a maximum score of 30/30. The average score achieved

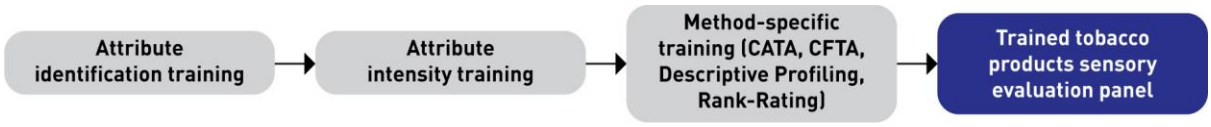
by the 109 individuals across the 30 samples in the eight assessment sessions was 85.2%. Scores ranged from a low of 43.3% (13/30) to a high of 100% (30/30). Average scores were comparable for each of the eight three-hour courses. All courses were presented by the same facilitators in an identical way, providing equality of opportunity for each and every candidate assessor.

The highest-scoring candidates were selected. Some individuals were eliminated from the shortlist based on them having displayed disruptive or non-conforming behaviour during the sensory aptitude sessions or providing a lower quality of answers in the initial questionnaire compared to those provided by other candidates. Forty selected candidate assessors were invited to attend an in-person independent assessment of their behavioural traits by a qualified Behavioural Psychologist. Personality profiles of individual candidate assessors were compared against a reference group of 3,915 people. A detailed report was prepared on each candidate assessor and reviewed. Taking into account all of the data collected about the candidates, 34 people were selected for training. The gender split of the pool was 73.5% female and 26.5% male; The average age of the assessors was 42.5 years, with a range of 23 – 68 years. The group of assessor trainees achieved an average score in the aptitude assessment tests of 94.8% (range 90 – 100%). This compares with an overall average for all external candidate assessors of 84.4% (range 43.3 – 100%) and an overall average of 77.0% for the group of external assessors who were not selected for training. The list of proposed candidate assessors and corresponding justifications for inclusion in the training programme were provided to the Commission and approved.

3.2 Training of the sensory panel

Each of the 34 trainee assessors who had been selected underwent ten days of core training followed by five days of method-specific training for a total of 15 days. The overview of assessor training and evaluation is depicted in Figure 2. The training program, which was compliant with the guidance contained in ISO 8586:2012 and ISO 5496:2006, was designed by an experienced trainer of professional sensory panels and delivered by that trainer, together with a qualified assistant.

Figure 2. Overview of assessor training and evaluation



3.2.1 Training design

Description of the aroma characteristics of test products by sensory assessors required the use of a defined set of descriptors. Where possible, such descriptors were anchored to reference materials allowing the meaning of the terms to be clearly and conveniently communicated to others via these reference materials. An initial list of descriptors was compiled through a combination of literature research and sensory evaluation of tobacco products by expert sensory assessors. In addition, the service listings of commercial laboratories which commonly support tobacco industry through analysis of flavour compounds in tobacco products were taken into account. With the active participation of trainee assessors, the list of attributes was further developed as the training developed. In this training a list of 51 odour descriptors was used (see Annex 2).

The assessors underwent several tests where they learned to recognise specific aromas in order to score them using a numerical scale. The training comprised of the following stages:

- Training in detection and recognition of attributes;
- Training in the rating of attribute intensity;
- Training in the use of specific sensory evaluation methods.

The following sections briefly summarize the tests performed.

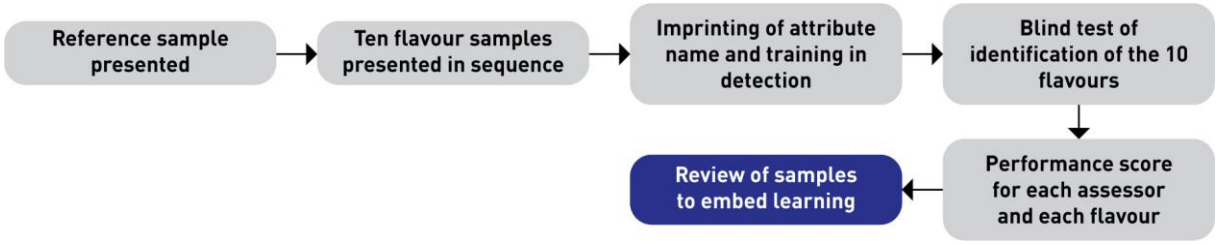
3.2.1.1 Aroma description test

Assessors were asked to provide a full description of the sample presented to them, drawing upon past experiences, including foods and beverages they have tasted, smell memories from home, work or travel, and any other suitable reference points. Descriptors generated by the group were collated on a flip chart, then shared and discussed with the group of assessors. The purpose of this test was to build experience, skill and community within the pool of assessors.

3.2.1.2 Aroma recognition test

A training session comprising four distinct stages as shown in Figure 3 was performed. The primary goal of this type of training session was to teach assessors to recognize a range of aroma compounds and use specific names given to them by a skilled facilitator, to identify them.

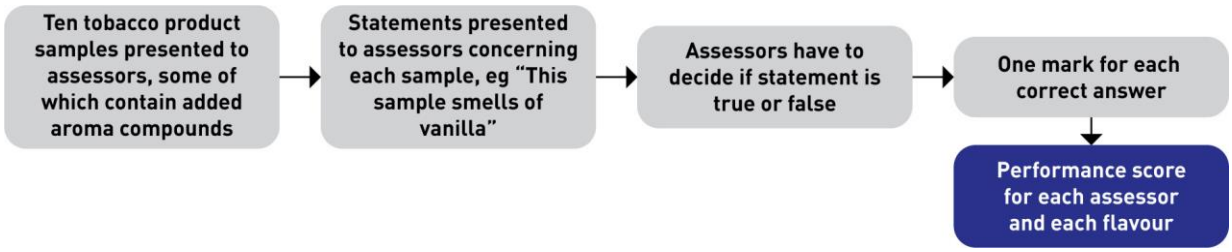
Figure 3. Overview of the aroma recognition training session



3.2.1.3 True/false test

As shown in Figure 4, assessors have to state if a statement like 'This sample smells of vanilla" is true or false for the sample presented. This type of test allowed assessors to consolidate the skills they learned in the aroma recognition tests.

Figure 4. Overview of true-false training session



3.2.2 Method specific training

In each method the assessors evaluated a set of samples, one by one, about which they have to answer specific questions. Several sensory tests were used in this method specific training procedure for the sensory panel:

- Rank-rating test: In this test, the assessors have to rank samples in order of the intensity of a specific odour that was added to the samples in different concentrations.
- CATA ('Check All That Apply') test: Assessors have to mark the presence/absence of the specific descriptors from the list of 51 odour descriptors, for each sample they receive.
- CFTA ('Check Five That Apply') test: Assessors are to name which five of the list of 51 odour descriptors are the most intense in the sample they receive.
- Descriptive profiling test: Assessors indicate the perceived intensity of each of the 51 odour descriptors for samples presented to them.

3.2.3 Assessment of sensory panel performance

Assessor performance was evaluated based on the scores achieved by the assessors in the training sessions. Performance scores attributed to each individual assessor were thus calculated based on their evaluations of approximately 240 individual aroma samples and more than 8,000 samples for the entire group of 34 assessors.

The results were analysed to allow the performance of the group as a whole, and the individual assessors to be compared. Statistical analyses, including Cochran's Q test, Analysis of Variance (ANOVA) and Principal Component Analysis (PCA) were performed using XLSTAT[20].

The performance of the group in tests comprising approximately 200 samples for which a specific session score was allocated to each assessor, averaged 97.3%, with a range of 88.7 – 99.5%. In large part, the scores confirm the efficacy of the recruitment, screening and selection techniques used to choose candidate assessors for training.

Each assessor underwent more than 75 hours of initial intensive sensory training and receive ongoing, routine maintenance and performance monitoring. Analysis of the data generated by the assessors demonstrated that both the group as a whole and the individual assessors were able to use the sensory techniques, they had learned to describe the aroma of tobacco products in both qualitative and quantitative terms.

[20] XLSTAT Sensory, www.xlstat.com

4 METHODOLOGY FOR THE TECHNICAL ASSESSMENT OF A TOBACCO PRODUCT

4.1 Establishing the list of reference products

4.1.1 Overview of sampling approach

4.1.1.1 Overview of sampling approach for cigarettes and roll-your-own tobacco

A random sampling among strata approach was used such as to capture a sample representative of products on the EU market, with regards to proper geographical distribution across EU member states (EU MS), for both leading and non-leading brands (oversampling of products with higher market sales based on data from Euromonitor) and broad distribution of tobacco product characteristics (tobacco part, leaf type, cure method, and additives).

Random sample selections were performed from the following strata:

- Stratum 1: Top market products in each EU MS market ("top market" sample), which included the top products per EU MS for boxed cigarettes and RYO tobacco per EU MS. The data source used for a market analysis for determining "top market" products was the Euromonitor Passport database[21]. This database, owned by Euromonitor International, contains the retail volumes and market share values of tobacco products in the EU. The rationale for oversampling the "top market" products is based on the fact that these products would have the most substantial population-based impact and would cover a more significant segment of the EU population. To verify this criterion, we evaluated the coverage of the top 10 boxed cigarette and RYO products from EU MS according to the 2016 Euromonitor data – which identified that on average ~70% of the total market share of the respective EU MS is covered by the top 10 selling products.

- Stratum 2: All products in the EU MS that are not among the top products in that market ("niche" or "non-top market" sample). Sampling within this stratum was performed using the entire list of products on each EU-MS market, using data collected from the EU Common Entry Gate (EU-CEG). This database is a comprehensive list, as, under the EU Tobacco Products Directive (2014/40/EU), manufacturers and importers of tobacco products are required to report information at the MS level on products which they intend to place on the market.

4.1.1.2 Overview of sampling approach for HTPs

For HTPs, the sampling approach involved the following steps:

- In late 2022, DG SANTE provided EU-CEG data which contained information on HTPs active in the EU market in 2021 along with corresponding sales data. This dataset was utilised to compile a list of HTPs for which active sales data for 2021 were available, leading to an identification of approximately 50 HTPs matching this criterion.
- In an effort to focus on products that are less likely to contain "characterising flavours other than that of tobacco", a manual search was conducted on the websites of the HTP companies included in the initial list. Any products that were evidently flavoured (such as those boasting tastes like berry, menthol, cool, watermelon and so forth) were removed. Furthermore, products that incorporated crushable flavour-containing capsules, which could alter the flavour of the product when employed, were also removed from the list. The aim of this procedure was to intentionally select a list of products that were more likely to be devoid of characterising flavours other than that of tobacco. This would facilitate the initial assessment of the HTP methodology, contributing to the creation of a reference space containing products with no or as little characterising flavours as possible for HTPs.

[21] Euromonitor International. Passport Data. https://go.euromonitor.com/passport.html. Published 2016. Accessed 5 June, 2020

4.1.2 Creating the final list of products

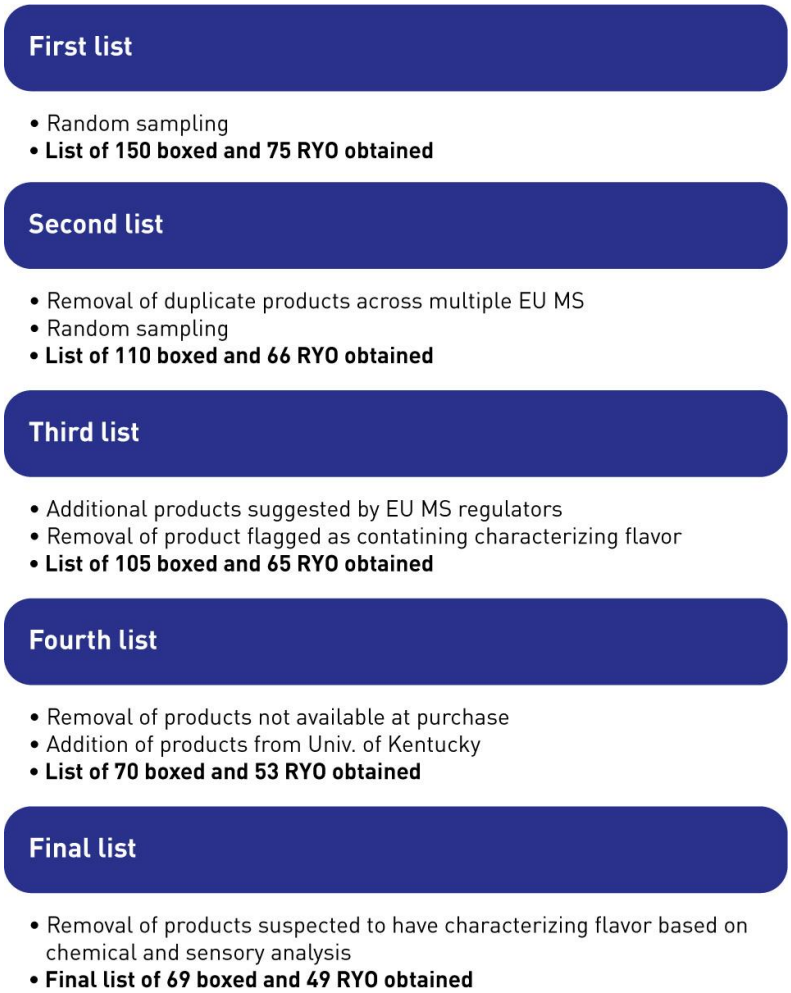
4.1.2.1 Creating the final list of test boxed cigarettes and roll-your-own tobacco for assessment

A final list of boxed cigarettes and roll-your-own products was created following the workflow depicted in Figure 5.

Briefly, a random sampling of boxed cigarettes and RYO tobacco products, representative of both top market and niche market products was performed to create a first list of products. Additional products from this list were oversampled for products to serve as a replenishment should a product have to be removed from the first list. A random sampling was performed to select 150 boxed products and 75 RYO products (Figure 9, Annex I).

Products flagged in EU-CEG as containing a characterising flavour by industry were removed from the first list, generating a second list containing 110 boxed products and 66 RYO products.

Figure 5. Overview of the process for creating the final list of products



Subsequently, EU MS regulators were asked to provide feedback on the national products within the second list. Through a questionnaire, EU MS regulators were asked whether there were any additional products in their EU MS, whose characteristics may not be represented by the proposed second list of products. These products were added to those already selected, so as not to interfere with the random sampling strategy. A third list of products was generated and samples represents products from EU MS, with an average of 6.2±1.4 products selected per EU MS, ranging from 4 to 9 products (see Table 2, Annex I).

Purchasing of products was completed with the assistance of the National Competent Authorities and other partners and shipped from each EU MS directly to the sensory testing facility via courier. Out of the 155 products of the third list, 36 were found to be no longer available on the market during the purchase process. Therefore, the purchase and shipping was completed for a total of 119 products (68 boxed and 51 RYO). Furthermore, 4 reference tobacco products (i.e. non-commercial, research-only tobacco products), two boxed cigarettes and two ground tobacco products, containing no additives were also obtained from the Centre for Tobacco Reference Products of the University of Kentucky. These products were included in the testing program, blind-coded together with routine test samples. Therefore, in total, 123 products were purchased for testing which constitute the fourth list of products (ranging from two to eight products per EU MS as per the random sampling plan).

TG tested the 123 products of the fourth list with sensory analysis (descriptive profiling) and chemical analysis, in order to determine the sensory and chemical characteristics. Some of the products were suspected to have characterising flavour based on chemical and sensory analysis and were removed from the fourth list in order to create a final list of products (The overview of the number of samples by EU MS is depicted in Table 3, Annex I).

4.1.2.2 Creating the final list of HTPs for assessment

The selection process for product assessment is presented in Figure 9, Annex I. In the first round, a pragmatic inclusion criterion was applied that involved choosing products based on specific brand marketing characteristics as described in section 4.1.1.2. To streamline the efficient purchase and timely shipment of products, products were purchased from 4 EU MS. Additional products were identified in subsequent rounds from several EU MSs in order to expand the geographical variability of the selected products. As a result of this purposive sampling, a total of 27 HTPs underwent sensory and chemical analyses, covering 24 EU MS markets.

An analysis undertaken by the EU-CEG indicated that the products selected via this procedure were available in a multitude of EU MS, under identical brand names. Minor variations were observed primarily in the composition of the filter and non-flavour-related compounds. However, these variations did not restrict the broad applicability of the findings to other EU MS where these products are sold.

4.1.3 Sensory assessment of selected reference products

Sensory analysis through the descriptive profiling of the selected products was conducted through the methodology of the assessment of a test tobacco described in section 3.2.1. For each product, the sensory method of descriptive profiling was conducted for the assessment of individual aroma attributes. Twelve trained sensory assessors evaluated each blind coded test sample in triplicate, rating the intensity of 51 individual odour attributes using a response scale of 0 – 10 (see Table 4 of Annex II).

Sensory assessors, given their extensive training in identifying the presence and intensity of specific attributes, could add new attributes not listed among those for tobacco, roll-your-own, or HTPs to their evaluations. Any such addition, marking both its presence and intensity, did not alter the procedure or initiate any further action, but was considered within the context of the specific product assessment. Intensity scores for each attribute, for each tobacco sample, were recorded by the assessors. Average individual attribute intensities and standard deviations were calculated for each test sample across all assessors. These calculations reflected the results assigned to the test sample by each assessor across all three replicates, amounting to 36 individual data points per test sample.

Samples with one or more attributes not originating from tobacco among the top 10 highest odour intensities (out of 51 core attributes for cigarettes and roll-your-own, and 54 attributes for HTPs), or with one or more attributes displaying average intensities that statistically differed from the reference products, were earmarked for further sensory testing through rank-rating.

4.1.4 Rank rating for the assessment of composite odour intensity

A rank rating test was used to assess the intensity of the overall composite aroma for samples that were flagged as possibly containing a characterising flavour, during descriptive profiling. During rank-rating, the composite odour of the sample is evaluated, rather than individual odour elements, reflecting how they are likely to be perceived by consumers. For example, a product may possess a composite aroma of *mint chocolate* comprising individual odour attributes of spearmint, peppermint, burnt sugar, vanilla and dark chocolate.

Sensory panellists were instructed to rate the intensity of such a specified composite aroma in each flagged sample using a scale of 0 – 10. For assessing if a tobacco product has a clearly noticeable flavour, the average intensity of the composite odour of the undiluted test sample is statistically compared with that of a reference product and with cut-off limits considered to represent a clearly noticeable odour intensity.

4.1.5 Principal Component Analysis (PCA) Analysis of reference products

Analysis of mean attribute intensity data for the potential reference products was performed through Principal Component Analysis (PCA), with the inclusion of 95% confidence ellipses on the sample plots to reveal the degree of similarity between replicate test results. The main objective of PCA is to summarize the information contained in the data set into a graphical representation of samples and variables by breaking it down into principal components. Two separate PCA representations were created, one for cigarettes and one for RYO tobacco. Confidence ellipses were based on data from three individual replicate analyses per product - coordinates of 95% confidence ellipses were defined using the function "coord.ellipse" in R (package: FactoMineR)[22,23]. PCA produces a map (space) in which the similarities and differences between products are depicted. PCA is carried out on the average (i.e. averaged over assessors in a sensory panel) intensity values for the set of individual odour attributes using XLSTAT.

For the final selected sample of boxed cigarette products, PC1 accounted for 25.2% of the variation in the data, with PC2 accounting for a further 9.2%. For the RYO tobacco products, PC1 accounted for 25.7% of the variation in the data, with PC2 accounting for a further 12.4%. For HTPs, using 54 attributes (51 core attributes plus three additional attributes: bacon, blackcurrant, eucalyptus), PC1 accounted for 34.7% of the variation of the data, with PC2 accounting for a further 18.2%.

A test of robustness of the multivariate spaces generated by PCA from average attribute data of replicate tests to assess *assessor response impact* was carried out. Simulated data sets of average intensities for each of the attributes were generated using the Monte Carlo approach. Standard deviations were calculated for each attribute and used to generate four alternative data sets for each reference space. PCA plots were made for each data set. A test of the robustness of the multivariate spaces generated by PCA from average attribute data of replicate tests to assess *product and sample impact* was also carried out. Simulated data sets of average attribute intensities for each odour were generated in which 10% of the products or samples (selected using a random function) were omitted from the data set. Products or samples were selected for elimination using a random function. PCA plots were made for each data set.

4.1.6 Chemical testing of reference products

Following the sensory assessment, a qualitative chemical analysis was performed. Each test product underwent chemical evaluation through headspace solid phase microextraction (SPME) analysis in combination with gas chromatography-mass spectrometry (GC-MS). Using this analytical approach,

[22] Le, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software. 25(1). pp. 1-18. http://www.jstatsoft.org/v25/i01/

[23] https://www.rdocumentation.org/packages/FactoMineR/versions/2.2/topics/coord.ellipse

the compounds in the volatile fraction of the test sample were identified using well-known MS libraries.

4.2 Overview of the technical assessment of a test tobacco product

What follows is a description of the procedure for when a test product goes through an in-depth assessment. The decision tree for the technical assessment of a tobacco product to support the decision of whether a tobacco product has a characterising flavour is shown below. The following sections explain each part of the decision tree (Figure 6).

4.2.1 Step 1. Sensory analysis of test sample through descriptive profiling

4.2.1.1 Step 1.1 Sample receipt, registration and storage

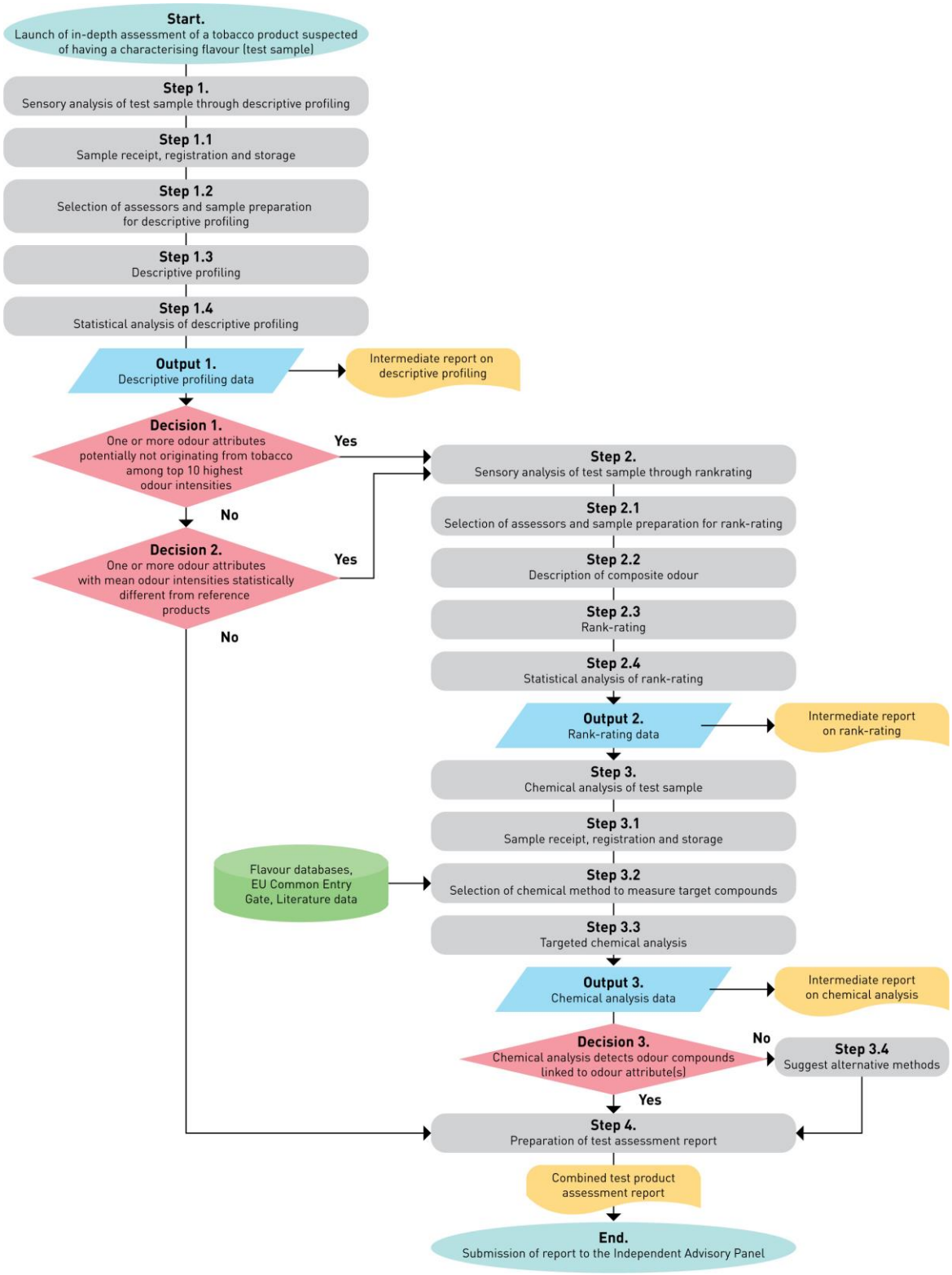
The sender provides details of the shipment and include a letter of authorization from the European Commission with each shipment, explaining the intended purpose of the shipment. If a shorter timeframe is required, necessary steps are taken by a member of the TG to arrange the pickup or purchase of the product within the respective EU MS.

For boxed cigarettes, a minimum of eight boxes of cigarettes (for the basis of 20 cigs/pack or more of necessary) should be purchased for testing. For RYO tobacco, a minimum of eight pouches (minimum of 30 g per pouch, with a minimum of three separate units) should be purchased for testing. For HTPs, a larger number of sticks are needed to compensate for the lower quantity of tobacco in HTPs when compared to cigarettes.

On receipt of the shipment at the sensory testing facility, details of the products are logged in the incoming sample schedule, the products are photographed, and the photographs and the other data are stored electronically.

All samples logged in the incoming sample are re-packaged and stored within 24 hours of their receipt. The sample boxed cigarettes, roll-your-own tobacco or HTPs, are placed in plastic-lined aluminium foil zip-lock bags and heat-sealed. Only products of a single type and brand are included in any one aluminium foil bag. The bag is transferred into the cold storage facility and stored at 5 – 7^{o}C until one day prior to testing.

Figure 6. Decision tree for the TG in-depth assessment of a tobacco product suspected of having a characterising flavour



4.2.1.2 Step 1.2 Selection of sensory assessors and sample preparation for descriptive profiling

For all sensory evaluation sessions, twelve assessors for each testing date are selected, and two assessors are allocated as reserves.

The products are removed from cold storage and attemperated to 20°C in their aluminium foil bags in a temperature-controlled environment for one to two hours. The first tobacco product is taken into the working area. The tobacco from the cigarettes is removed by cutting each cigarette longitudinally with a sharp-sanitized knife. The required quantity of tobacco will be transferred onto a disposable plastic (odourless) weighing boat. The tobacco is weighed using a calibrated digital balance, and the weight is recorded in the sample production record. The required quantity of tobacco will be transferred into the electric blender, and the sample is ground for three cycles of 5-seconds each. The electric blenders used for the preparation of tobacco samples are used exclusively for this purpose. The ground tobacco is placed into a labelled zip-lock plastic bag. When all samples have been ground and transferred to plastic bags, a semi-automatic pharmaceutical-grade capsule filler is used to fill gelatine capsules (size 00) with 0.35 g of ground tobacco.

Ninety capsules of tobacco, sufficient for each test sample to be evaluated in triplicate, is prepared. Confirmation that each capsule contains 0.35 g $\pm$ 0.02 g of ground tobacco is achieved by performing check weights on 10 capsules from every batch. Two capsules of tobacco are transferred into individual clean screw-cap glass vials (20 ml capacity). Each glass vial, therefore, contains 0.7 g of ground tobacco. For HTPs, rather than using a fixed weight of 0.7g of tobacco material in each sample vial, a quantity of each product in the tube that occupies the same volume as 0.7g of a typical sample of cigarette tobacco is applied. This is necessary as some of the tobacco material in HTPs is of very low density and, thus, occupies a considerably larger space in the sample tubes compared to cigarette tobacco. Attention is given to keep the volumes of test and decoy samples in the glass vials constant despite potentially different bulk densities.

Forty-two vials of each test sample are required for triplicate testing (36 samples plus contingency). A clean plug of cotton wool (0.25 g) is inserted into each vial, and each vial is sealed with a foil-lined plastic screw cap. The surface of each glass vial is cleaned with an odourless alcohol wipe. A three-digit blinding code label is applied to each vial (same code for each individual replicate for a single test sample – three codes for three replicate test samples for a single batch). Sample vials are prepared 16 – 20 hours in advance of sensory assessment and stored at 19-20°C in the dark in clean cardboard sample boxes prior to use.

In addition to the test product, a range of decoy and reference products are also selected for parallel assessment. Decoy products can include unflavoured tobacco samples, flavoured tobacco samples and previously-tested samples. Reference products include single brands of boxed cigarettes, RYO tobacco or HTPs from a single lot or well-characterised reference cigarettes or powdered tobacco from the University of Kentucky. Both cigarettes and RYO tobacco are evaluated on the same testing day and within a single testing session. Decoy samples may also comprise previously tested sample vials and tobacco samples to which flavour substances have been added.
The quantity of tobacco required for testing is as follows:

- Test sample (boxed cigarette) – triplicate determination – 4 boxes of 20 cigarettes
- Test sample (roll-your-own) – triplicate determination – 2 pouches of 30 g each
- Test sample (HTPs) – triplicate determination - 8 boxes of 20 sticks
- Reference samples (boxed cigarettes) – single determination – 1 box of 20 cigarettes
- Reference samples (roll-your-own tobacco) – single determination – 1 pouch of 30 g
- Reference sample (HTPs) – single determination – 2 boxes of 20 sticks
- Decoy samples (boxed cigarettes) – single determination – 1 box of 20 cigarettes
- Decoy samples (roll-your-own tobacco) – single determination – 1 pouch of 30 g

4.2.1.3 Step 1.3 Descriptive profiling

Twelve trained sensory assessors evaluate each blind coded test sample in triplicate, rating the intensity of individual odour attributes using a numeric response scale (0 -10, whole numbers only, where 0 = absent).
The test samples are complemented with reference and decoy samples as described above and are presented to the assessors in a defined sequence, which differs for each assessor.

Should any sensory assessor identify any unexpected odour, not among the attributes that they are trained on, they have the choice to add this new attribute to their response form, noting both the presence and intensity of such additional attributes. Each assessor completes the evaluation of their first sample prior to commencing their assessment of their second and subsequent samples. To minimize possible expectation bias in future evaluations, no feedback on sample evaluations is provided to assessors following descriptive profiling of test samples.

The expert sensory panel that carries out the technical sensory assessments of test products has been trained using a fixed vocabulary – directly linked to chemical substances and individual odour attributes- rather than a vocabulary derived from, and associated with, a particular sensory panel. Specific labels are used to identify specific chemicals. Thus, assessors trained to evaluate the flavour of tobacco products are able to identify and rate the intensity of individual flavour compounds (such as β-ionone), rather than refer to them by means of collective terms that may be sensory panel specific (such as floral).

Assessors must not bring personal effects, food or drink into the Sensory Evaluation suite. They must also refrain from wearing scented products. Instructions concerning the sequence in which the test samples evaluated are provided by the software. Indicators are provided in the software to direct the sensory assessor in the sample evaluation process. A 10-minutes break is provided for assessors between each evaluation session. Assessors are not permitted to discuss the samples or attributes associated with samples at any time.

4.2.1.4 Step 1.4 Statistical analysis of descriptive profiling

Mean individual attribute intensities and standard deviations are calculated for each test sample. This reflects the results for the test sample allocated by each of the 12 assessors for all three replicates, representing 36 individual data points per test sample. Results are shown by a graphical representation of the attribute intensities in form of a ranked list from highest to lowest, including the individual mean score of every attribute and their 95% confidence interval (CI) (error bars) together with the mean, lower limit and upper limit attribute scores for the reference products of that category (RYO or boxed, respectively). The results are graphically presented in the form of a spider diagram.

Individual assessor responses from all samples evaluated on a single day of testing are subject to Analysis of Variance (ANOVA) to confirm the validity of the test results and the variance between assessors.

An Analysis of Variance (ANOVA) is carried out to calculate the Least Significant Difference (LSD) between attribute intensities of the test product and reference products for each attribute. When the difference between two products is higher than the LSD, those products can be considered statistically different in terms of that attribute. The LSD is displayed on the bar chart as an error bar (95% confidence interval).

The attributes in this sorted list are checked for the presence of those considered to potentially originate from a source other than tobacco. The attributes that potentially originate from tobacco or a source other than tobacco are presented in Annex II and described in Table 1. The lists of attributes are only used within the context of descriptive profiling for determining whether the test product goes on for additional testing and should not be interpreted for the purpose of applying the TPD.

Should a product have within the top 10 ranked attributes an attribute that potentially originates from a source other than tobacco, the product undergoes additional sensory assessment through rank

rating. This top 10 cut off is used as a screening process to minimize the risk of false-negative results arising from descriptive profiling (a false-negative result would arise if a sample that actually had a characterising flavour was classified as a negative sample by the test result), while accepting a reasonable incidence of false-positive results (the case in which a sample without a characterising flavour would be classified as a positive sample). This top 10 screening test is used to avoid the unnecessary burden of having every tested product undergo both descriptive profiling and rank rating. However, all products are subject to additional testing to support the decision on whether a tobacco product has a characterising flavour other than that of tobacco.

The test sample is projected onto the multidimensional Principal Component Analysis (PCA) product space of the reference products, boxed cigarettes, RYO, or HTP, respectively. PCA is carried out on the average intensity values for individual odour attributes using XLSTAT and is performed to provide supplementary evidence relating to the sensory characteristics of test products.

Table 1. Attributes and their potential origin for the purpose of descriptive profiling

Odour attributes that may potentially originate from tobacco include:
• Odour attributes such as green pepper, potato skin, citronella and cedar, which may be derived directly from tobacco plant metabolism; • Odour attributes such as black tea, rotted dry wood, violet and saffron, which may be derived from the chemical breakdown of carotenoid pigments in tobacco leaves during curing; • Odour attributes such as cardboard, cucumber, freshly cut grass and hay, which may be derived from the oxidation of lipids in tobacco leaves during curing; • Odour attributes such as cheese, which may be derived from hydrolysis of lipids in tobacco leaves during curing; • Odour attributes, such as sweetcorn, which may be derived from the breakdown of amino acids during drying of tobacco; • Odour attributes, such as vinegar which may be derived from the action of microorganisms during tobacco processing; • Odour attributes, such as smoky and burnt coffee, which may originate in the smoke generated in materials other than tobacco which may be used to dry “fire-cured” tobacco; • Odour attributes, such as dried leaves, prune and raisin, which may be of uncertain origin but found to occur at noticeable levels in all commercial tobacco products tested.
Odour attributes that may potentially originate from a source other than tobacco include:
• Fruity odour attributes, such as artificial apple, artificial cherry, banana, coconut, grape, green banana, mango, orange-limonene, peach, raspberry or strawberry milkshake; • Mint-like odour attributes, such as menthol, menthone, methyl acetate, and spearmint; • Sweet attributes, such as burnt sugar, dark chocolate, honey, maple syrup, and molasses; • Confectionary-like odour attributes, such as butter, candy floss, marzipan, and vanilla; • Spicy odour attributes, such as aniseed, cardamom, cinnamon, clove, and coriander seed; • Floral odour attributes, such as lavender and rose.

During the assessment of the HTPs, additional odours other than those typical of tobacco, which are not represented within the above-mentioned attributes, have been detected in some of the HTP samples tested to date, including blackcurrant, eucalyptus and bacon.

4.2.2 Step 2. Sensory analysis of test sample through rank-rating

4.2.2.1 Step 2.1 Selection of assessors and sample preparation for rank-rating

Two types of tobacco samples are needed for this test: A reference tobacco product and the test sample suspected of possessing a characterising odour based on descriptive profiling results. The details of the two tobacco products are confirmed in the incoming sample schedule.

The quantity of tobacco required for testing is equivalent as follows:

- Test sample (boxed) – 2 boxes of 20 cigarettes
- Reference sample (boxed) – 2 boxes of 20 cigarettes
- Test sample (RYO) – 1 pouch of 30 g each
- Reference sample (RYO) – 1 pouch of 30 g each
- Test sample (HTPs) – 4 boxes of 20 sticks
- Reference sample (HTPs) – 4 boxes of 20 sticks

The blends are prepared per the same procedure used for preparing samples for descriptive profiling described in section 4.2.1.2. For rank-rating, thirty capsules of tobacco for fifteen vials for each test sample to be evaluated in triplicate is prepared.

When the samples have been prepared, a series of mixtures of the two samples is prepared. To avoid expectation bias among sensory assessors, the ratio of reference and test samples vary from one test to the next. Typically, assessors are presented with eight test samples, including the test sample, a known reference tobacco product with no characterising odour, and several mixtures containing each of the two samples in different proportions. An example of a set of samples is:

- Reference tobacco (100%) – test sample (0%) – two replicates
- Reference tobacco (75%) – test sample (25%) – one replicate
- Reference tobacco (50%) – test sample (50%) – two replicates
- Reference tobacco (25%) – test sample (75%) – one replicate
- Reference tobacco (0%) – test sample (100%) – two replicates

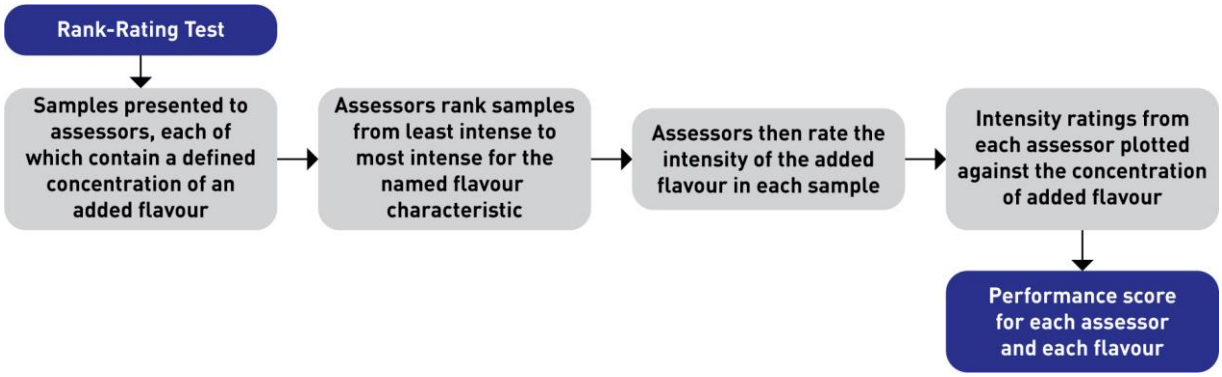
4.2.2.2 Step 2.2 Description of the composite odour

The construct odour to be used is agreed with the sensory panel in advance of the evaluation. Prior to commencing the evaluation, the attribute to be evaluated is revealed to the assessors. This will mostly be a composite comprising the combined odour effect of several different odour compounds. In most cases, the composite odour term to be used in such evaluations will be obtained from a list. However, in cases in which the match cannot be found for the odour characteristics detected in the sample, the term to be used will be agreed with the sensory panel by consensus in advance of the rank-rating evaluation.

4.2.2.3 Step 2.3 Rank-rating

For evaluation by rank-rating, samples are assessed in two stages. In the first stage, the relative intensity of the composite odour of interest is determined for each sample and indicated by the assessor through physically moving the samples in front of them. In the second and final stage of the assessment, evaluators must rate the intensity of the composite attribute in each of the ranked samples using a scale of 0 - 10.

Figure 7. Rank-rating test flowchart



4.2.2.4 Step 2.4 Statistical analysis of rank-rating

Because the term *clearly noticeable* has no specific meaning in sensory science, we define the following criteria for that:

1. It is distinctive from and contrasting to that of the odour characteristics of the tobacco in which it is present.
2. It is more intense than that which would be regarded as “noticeable”.

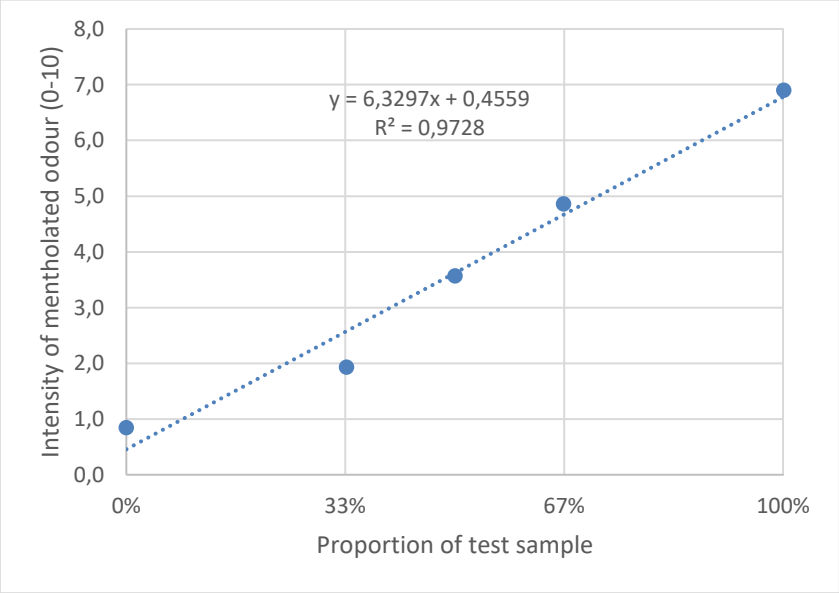
A line of best fit is calculated for the relationship between the proportion of test sample included in the tobacco sample (0 - 100%) and the odour intensity perceived by each assessor for each sample. The coefficient of determination (R2) is calculated for the proportion of test sample included in the evaluated tobacco sample and the average corrected odour intensity perceived by the panel for all samples and replicates, which is used as an indication that the test sample may possess a clearly noticeable flavour.

In addition to the coefficient of determination, the difference in the cut-off values between the two types of products (test vs. reference) is intended to reflect the ease with which the characterising flavour can be noticed against the overall odour of the tobacco product. The intensity of the construct odour in rank-rating rated >2.5 for RYO and HTPs, and >1.8 for cigarettes (0-10 scale) and is regarded also as an indication that the test sample may possess a clearly noticeable flavour [24,25].

More specifically, assessment of the reference tobacco products indicated that RYO and HTP products, in general, had higher intensity scores than boxed cigarettes. A permutation test is used to compare the mean intensity of the characterising odour in the test product with the cut-off limit of 1.8 (for boxed cigarettes) or 2.5 (for RYO tobacco and HTPs) and with the mean intensity of the characterising odour in the control product.

Rank-rating results are displayed in a scatter plot together with a linear regression of the data, indicating the overall odour intensity for the specific flavour (on a scale of 0 – 10) including the name of the odour and the potential outcome of the rank rating and the odour threshold number achieved by the sample (Figure 8).

Figure 8. Example rank-rating plot for evaluating the composite odour “mentholated” of a test product



[24] Odour threshold numbers are closely aligned with the statutory action limits for potable water in the UK water industry

[25] The Determination of Taste and Odour in Drinking Waters – Methods for the Examination of Waters and Associated Materials, Environment Agency, 2014

4.2.3 Step 3. Chemical analysis of the test sample

4.2.3.1 Step 3.1 Sample receipt, registration and storage

Upon their arrival, samples are placed in a refrigerator until the day before their analysis. Products are removed from the refrigerator 1-2 hours before being prepared to be placed into vials in order to reach ambient temperatures. All products remain in the plastic-lined aluminium foil zip-lock bags until actual analysis.

4.2.3.2 Step 3.2 Selection of chemical method to measure target compounds

So as to identify the compounds with a potential odour, the odour description is assessed with the use of the following databases, as necessary:

- Flavours databases may be used to identify chemical additives that are characteristic for a certain flavour, including: (1) The Leffingwell Flavour Database[26] which provides information relevant to the Flavour, Food, Beverage & Tobacco industries, extensive information on tobacco flavouring, including over 4000 descriptions of flavouring materials and additives and (2) The Good Scents Company Information System[27] which provides information for the Flavour, Food and Fragrance Industry.
- EU-Common Entry-Gate (EU-CEG) data reported by the tobacco industry may be used to list the chemical compounds in the test product. This list of compounds is reported using their CAS number and this list will be assessed for the existence of the substance(s) that could lead to the flavour observed through sensory analysis.
- If the evidence available through the above data sources is inconclusive, then a full scan analysis will be performed through the use of headspace solid-phase microextraction gas chromatography-mass spectrometry (HS-SPME-GCMS). During this analysis, the compounds in the test sample are detected, and chromatograms areas are exported. A chemical profile of the test product will be created, that can be used for comparison against other tobacco products. These would include products with and without a characterising odour. The individual peaks of the chromatogram, which reflect compounds in the test product, will be identified through the use of digital mass spectral libraries including FFNSC (Mass spectra of Flavors and Fragrances of Natural and Synthetic Compounds); NIST (National Institute of Standards and Technology); WILEY; and PMW (Pfleger, Maurer, Weber).

The matching compounds identified via the above data sources would be subsequently assessed through a “fit for purpose” chemical analysis. In the case of targeted analysis, the rationale for the choice of the specific chemical assessment, as well as method performance characteristics will be brought forward using internationally accepted guidelines.

4.2.3.3 Step 3.3 Targeted chemical analysis

Each tobacco sample is analysed in triplicate to ensure precision, while the respective chromatograms will be produced. The chromatograms will provide the respective peaks for the assessment of the compounds. The identified chemical compound or compounds that match the odour attributes also identified by the sensory assessment will undergo quantification, and the concentration of the compound will be measured. The validation of the applied method will be performed based on elementary analytical parameters. Calibration will be performed by internal standardisation. The applied method will be validated for Linearity; Limits of detection (LOD); Limits of quantification (LOQ); Precision; Recovery and Carryover effect.

[26] Leffingwell & Associates. Flavor-Base 9 - Tobacco Version. 2012. http://www.leffingwell.com/products.htm

[27] http://www.thegoodscentscompany.com/

4.2.4 Step 4. Preparation of test assessment report

Where IAP considers input from the technical group necessary for the purposes of providing an opinion, IAP shall have in forming its opinion regard to the information and data provided by the technical group in line with the methodology presented in this report.

For this purpose, results of all assessments performed with a test product should be merged into an overall “test assessment report” based on the outcome sensory analyses and complemented by the results of the chemical analysis, as necessary.

5 ANNEX I

Figure 9. Flowchart of a sampling plan for a high-level list of tobacco reference products

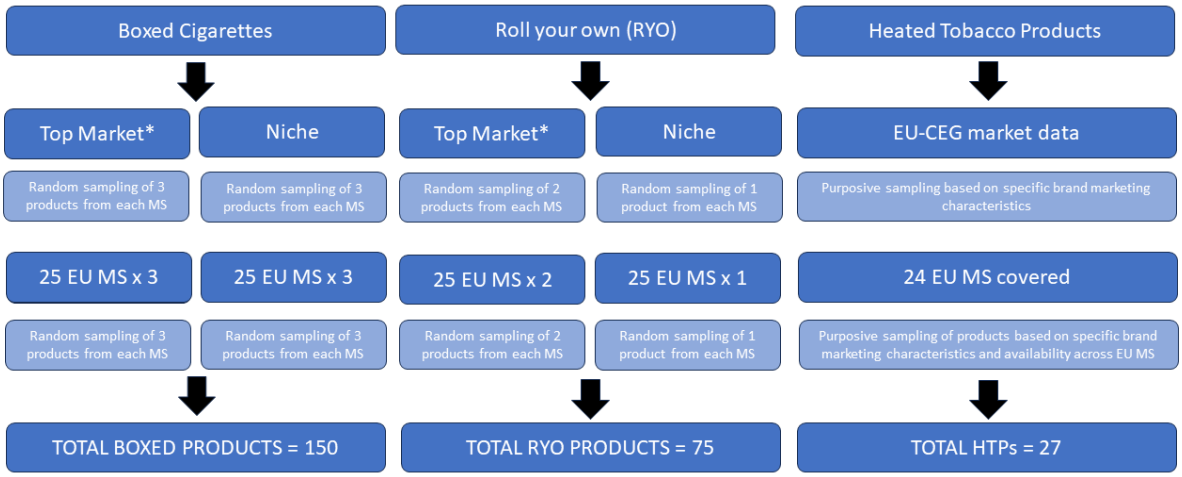


Table 2. Third list: number of selected boxed, RYO and HTP products by strata (top market and niche)

Strata	Number of Products
Boxed	**93**
Boxed Niche	42
Boxed Top 10	51
RYO	**62**
RYO Niche	24
RYO Top 6	38
HTP	**27**
Grand Total	**182**

Table 3. Final list of purchased products by EU MS

EU MS	Number of Boxed Products	Number of RYO Products	Number of HTP Products	Total
AT	1	3	0	4
BE	3	2	0	5
BG	3	2	0	5
HR	2	2	0	4
CZ	3	2	0	5
DK	2	3	3	8
EE	4	1	0	5
FI	2	3	0	5
FR	2	3	0	5
DE	5	3	0	8
GB	2	3	0	5
GR	3	3	19	25
HU	2	3	0	5
IE	4	3	0	7
IT	2	1	3	6
LV	2	2	0	4
LT	1	0	0	1
PL	2	1	0	3
PT	3	1	2	6
RO	5	1	0	6
SK	3	1	0	4
SI	5	0	0	5
ES	1	1	0	2
SE	2	2	0	4
NL	3	1	0	4
Univ. Kentucky	2	2	0	4
Grand Total	**69**	**49**	**27**	**145**

Sampling approach (see 4.1.1 for more details)

From the 123 products of the "fourth list", a total of 118 products (69 boxed cigarette and 49 RYO tobacco products) were confirmed by sensory analysis as not possessing one or more characterising flavours.

From the remaining five products, two products were removed because they contained a clearly noticeable characterising flavour(s). The first was a boxed cigarette product marketed in the EU as a mentholated product, and therefore should not a priori have been considered as a candidate reference product for testing. However, in the EU CEG database the product had been declared by the tobacco manufacturer/importer as not having a characterising flavour. Testing confirmed that the sample possessed a characteristic *mentholated* odour, which was clearly noticeable. The second were a RYO product because it had a characterising *apple juice* odour other than that of tobacco as indicated by sensory testing and confirmed by chemical testing.

Other three products were decided as having a characterising flavour for the attribute *smoky*. Given the distinct sensory characteristics of these products, in both the descriptive profiling and PCA analysis - and in the absence of information on the chemical compounds found in the samples that contribute to a smoky aroma - these products have been included among the reference products on the basis

that they belong to a specific sub-group with characterising *smoky* flavours, but are to be used only when a test product has a specific *smoky* aroma.

Concerning HTPs, an initial market and EU-CEG analysis was performed on 27 MSs to identify candidate products. To streamline the efficient purchase and timely shipment of products, products were purchased from 4 EU MS, with a market outreach across 24 EU MS.

6 ANNEX II

Table 4. List of 51 individual aroma attributes and compounds used in descriptive profiling of a test product, no. 52 to 54 added specifically for HTP.

No	Attribute	Descriptors	Aroma chemical
1	Aniseed	Fennel, ouzo, liquorice	*trans*-Anethole
2	Artificial apple	Synthetic apple, red apple	Ethyl hexanoate
3	Artificial cherry	Synthetic cherry, cherryade, scented candle, air freshener	Benzyl acetate
4	Banana	Fruity, pear drops	Isoamyl acetate
5	Black tea	Tobacco, blackcurrant, red wine	β-Damascenone
6	Burnt coffee	Instant coffee, coffee liqueur	Furfurylthiol
7	Burnt sugar	Caramel, caramelized strawberry	Furaneol
8	Butter	Buttermilk, warm milk, movie popcorn	Diacetyl
9	Candy floss	Cotton candy	Ethyl maltol
10	Cardamom	Cardamom pod, Ikea pine	Terpinyl acetate
11	Cardboard	Dry paper, dry cardboard	*trans*-2-Nonenal
12	Cedar	Dry wood, cigar box	β-Caryophyllene
13	Cheese	Limburger cheese, human sweat, old hops	Isovaleric acid
14	Cinnamon	Cinnamon toast, eggnog, hot cross bun	*trans*-Cinnamaldehyde
15	Citronella	Citronella candle, lemon verbena	β-Citronellol
16	Clove	Clove oil, spicy	Eugenol
17	Coconut	Shampoo, suntan lotion, Malibu rum	γ-Nonalactone
18	Coriander seed	Orange peel, tangerine, Belgian Wit beer	Linalool
19	Cucumber	Cucumber skin, watermelon	*trans,cis*-2,6-Nonadienal
20	Dark chocolate	High cocoa solids chocolate, burnt, ashy	2,3,5-Trimethylpyrazine
21	Dried leaves	Forrest floor, winter leaves	In development
22	Freshly cut grass	Grassy, cut leaf, green	*cis*-3-Hexenol
23	Grape	Grape juice, wine	Methyl anthranilate
24	Green banana	Plantain, under-ripe banana, banana skin	*cis*-3-hexenyl acetate
25	Green pepper	Earthy, bell pepper, green bean	2-Isobutyl-3-methoxypyrazine
26	Hay	Dry grass, late summer hay	In development
27	Honey	Fragrant, mead, rose	Phenylethyl acetate
28	Lavender	Flora, fragrant, Earl Grey tea	Linalyl acetate
29	Mango	Over-ripe mango, tinned pineapple	Ethyl butyrate
30	Marzipan	Christmas cake, almond, cherry	Benzaldehyde
31	Maple syrup	Fudge, tablet	Maple lactone
32	Menthol	Breath freshener, mentholated sweets, peppermint	Menthol
33	Menthone	Breath freshener, mentholated sweets, mint	Menthone
34	Menthyl acetate	Mint, damp forest floor	Menthyl acetate
35	Molasses	Treacle, liquorice	In development
36	Orange-limonene	Fresh orange juice, Fanta orange	D-Limonene
37	Peach	Peach, peach skin, crayon	γ-Undecalactone
38	Potato skin	Earthy, soil, bag of potatoes	2-Isopropyl-3-methoxypyrazine
39	Prune	Dried fruit, prune juice	In development
40	Raisin	Dried fruit, sultanas	In development
41	Raspberry	Raspberry juice, raspberry jam	Raspberry ketone
42	Rose	Floral, rose petals, perfume	Geraniol
43	Rotted dry wood	Decaying tree branches, hay, sweet tobacco	β-Cyclocitral
44	Saffron	Fragrant, sweet, oriental	Safranal

45	Smoky	Smoked cheese, smoked ham, smoked fish	Guaiacol
46	Spearmint	Mint, toothpaste	L-Carvone
47	Strawberry milkshake	Strawberry, apple juice, cider	Ethyl-2-methylbutyrate
48	Sweetcorn	Tinned tomatoes, strawberry jam, seaweed, truffle	Dimethyl sulphide
49	Vanilla	Ice cream, custard, barrel-aged wine	Vanillin
50	Vinegar	Vinegar, pungent	Acetic acid
51	Violet	Violet, freesia, rose water, Turkish delight	β-Ionone
52	Eucalyptus	Basil, oregano, sage, 'Dr Pepper', 'Deep Heat', tea tree oil	1,8-cineole
53	Bacon	Smoked bacon, smoked cheese, creosote, pork scratchings	4-ethyl guaiacol or 4-methyl guaiacol
54	Blackcurrant	Blackcurrant juice, blackcurrant lozenges, 'Cherry Tunes', strawberry liquorice	blackcurrant thiol, 4-mercapto-4-methylpentanone or other sulphur compounds

7 ANNEX III

Figure 10a. Example of spider diagram for cigarettes showing individual attribute intensities.

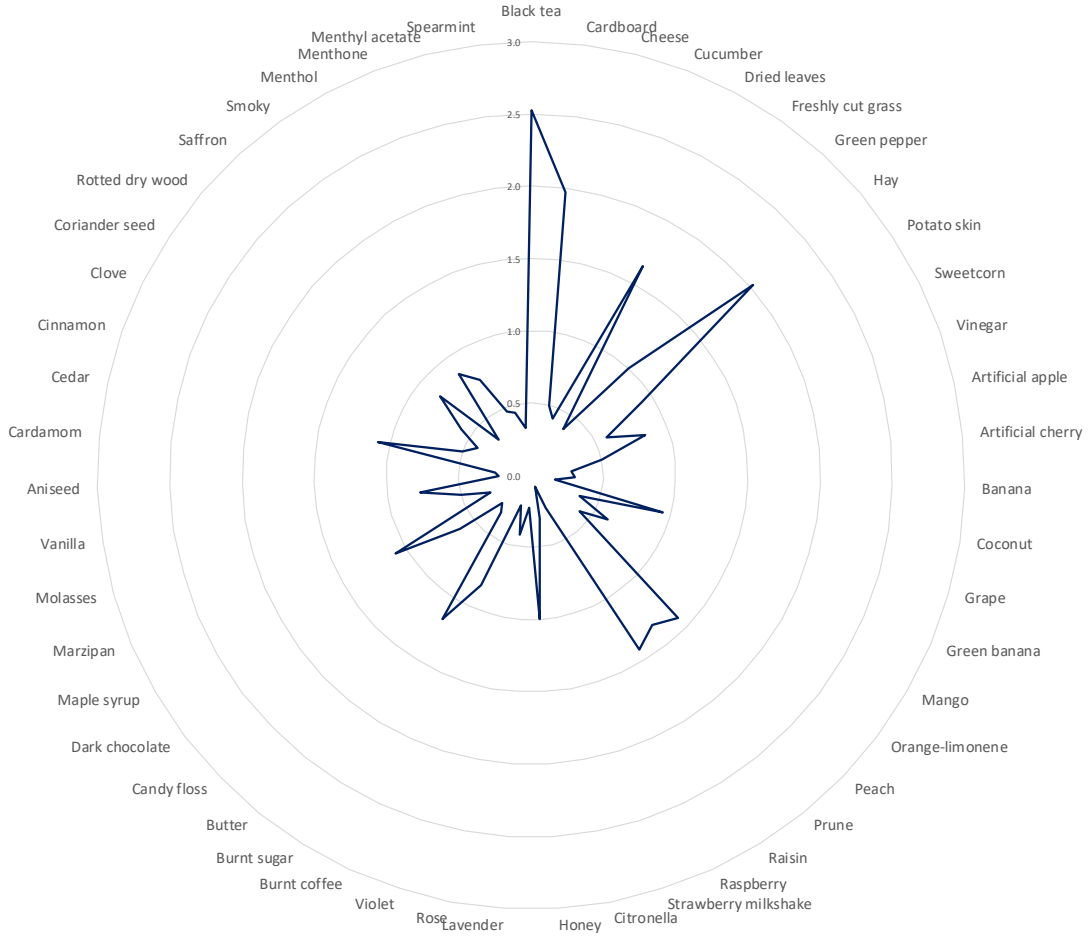


Figure 11b. Example of spider diagram for RYO showing individual attribute intensities.

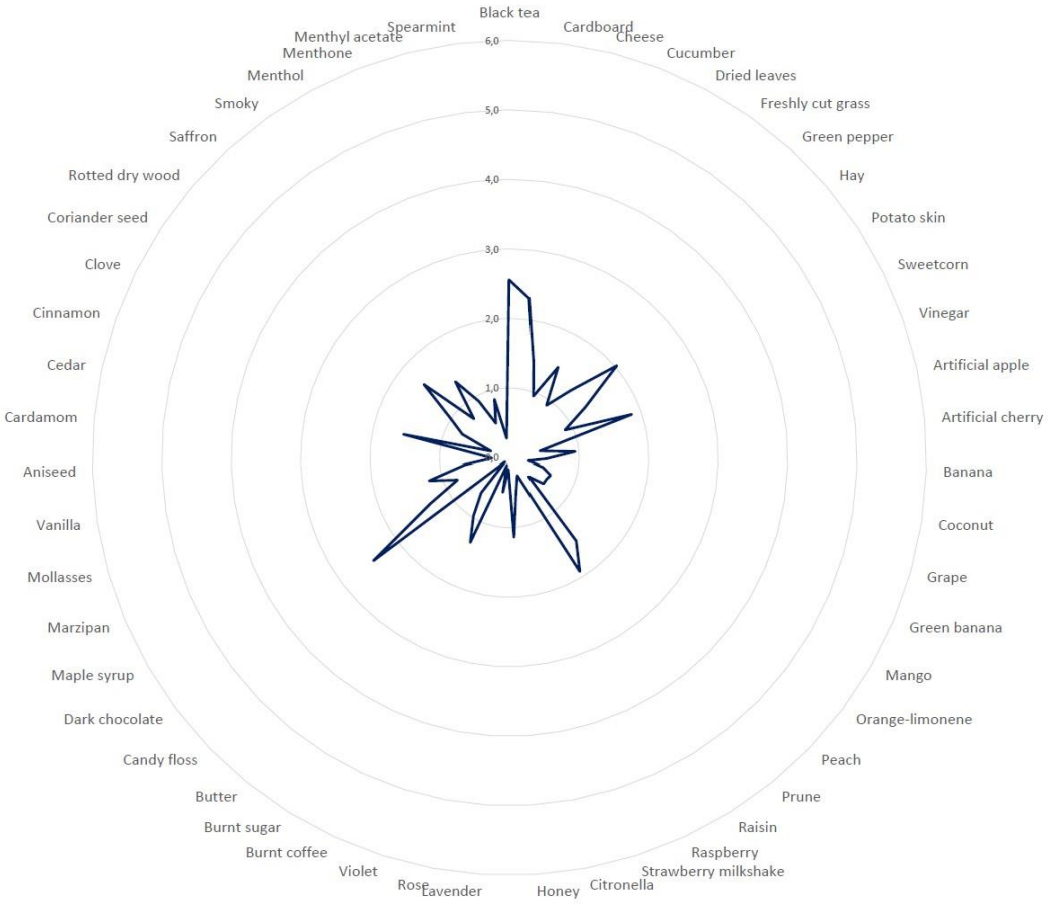


Figure 12c. Example of spider diagram for HTPs showing individual attribute intensities.