**UN-BRANDED PRODUCT TEST (UPT)**

**April 2013 v1**

LEGAL NOTICE

*ASU30*

*ASU30 and ATUU30 are acronyms for Adult Smoker Under 30 and Adult Tobacco User Under 30. The term ‘adult’ is defined by local law, but shall in no circumstance refer to any person under the age of 18. Likewise, consistent with BAT’s International Marketing Standards, terms such as ‘consumer’, ‘target consumer’, ‘smoker’, or ‘target audience’ refer only to smokers whom are adults.*

LOCAL LEGISLATION / REGULATIONS

*The material contained in this document may present executions and themes that are not legally permissible or acceptable in some markets. Likewise, because tobacco regulation and the company’s own views on socially responsible marketing change over time, previously approved activities may no longer be permissible. Therefore, you must obtain local legal and CORA approval before proceeding with any activity in your market.*

*COPYRIGHT AND CONFIDENTIALITY*

*© British-American Tobacco (Holdings) Limited 2013. All rights reserved. No part of this document may be reproduced in any form or any means without the prior written consent of British-American Tobacco (Holdings) Limited (˝BAT˝).*

*This document is proprietary to BAT and is provided to employees of the BAT Group and certain trusted contractors working for a member of the BAT Group. It is disclosed solely for use by employees in the course of their employment or by contractors for the purpose of facilitating the provision of services by that contractor to the employing BAT company and for no other purposes. Unauthorised possession or use of this material or disclosure of the proprietary information without the prior written consent of BAT may result in legal action.*

*THIRD PARTY TRADEMARKS*

*“Trademarks, logo designs and/or brand names featured in this material and not owned by British American Tobacco PLC (or one of its group companies) are owned by the relevant proprietor of the particular trademark and/or brand name in question, and are referred to in this document for informational and/or illustrative purposes only.”*

Contents

[1. Overview Un-branded Product Test 3](#_Toc353355274)

[2. How to setup un-branded product test? 5](#_Toc353355275)

[2.1 Research Design: 5](#_Toc353355276)

[2.2 Design Scenarios: 8](#_Toc353355277)

[2.3 Sample size and type: 13](#_Toc353355278)

[**2.4** **Rules around target group setting:** 14](#_Toc353355279)

[**2.5** **Stimulus materials and requirements:** 15](#_Toc353355280)

[2.6 Product masking: 16](#_Toc353355281)

[2.7 Product coding: 16](#_Toc353355282)

[2.8 Product handling 18](#_Toc353355283)

[3. Interview and questionnaire flow 19](#_Toc353355284)

[3.1 Visit and Interview flow: 19](#_Toc353355285)

[3.2 Questionnaire flow: 21](#_Toc353355286)

[3.3 SAS+ attributes – local translation and validation: 22](#_Toc353355287)

[3.4 Questionnaire length: 23](#_Toc353355288)

[4. Key Metrics and Analytics 24](#_Toc353355289)

[4.1 Key outcome measures: 24](#_Toc353355290)

[4.1.1 Product Performance: 24](#_Toc353355291)

[4.1.2 Consumer pathway to drive overall acceptability: 26](#_Toc353355292)

[4.1.3 Product attribute space map: 26](#_Toc353355293)

[4.1.4 Penalty analysis: 27](#_Toc353355294)

[4.1.5 Comparison against Normative Database: 27](#_Toc353355295)

[4.1.6 BAT Product Normative Database (C-PSI): 28](#_Toc353355296)

[4.2 ACTION STANDARD SETTING: 29](#_Toc353355297)

[5. Agency for Un-branded Product Test 31](#_Toc353355298)

# Overview Un-branded Product Test

Traditionally BAT has done most of the product testing in the un-branded format in order to understand the actual product performance without the impact of branding. With the new focus on product and developing differentiated and innovative products – BAT has revamped the product testing toolkit to include a wider range of tools both for un-branded and also branded product testing. The new toolkit will ensure that we have a comprehensive set of methods to do both strategic and tactical product testing for all the different types of products available.

The *Un-branded Product Test (UPT)* is an upgraded version of the previous CPT/MASQ/PSM methodology that BAT has used for the past 20 years. UPT has a rich set of Sensory Attributes specifically designed to suit a wide range of product types and it incorporates some new analytics to help us understand the performance of the differentiated and innovative products.

UPT replaces the previous CPT and PSM methods. The earlier MASQ testing is replaced by a Branded Product Test (BPT) and specifically the Brand Monitor.

In the new product testing toolkit the UPT will play an important role in testing the pure product performance and helping us to understand the true impact that branding and tactile elements have on consumers taste perception.

One of the most exciting features of the new UPT is that it links directly to the Product Norms Database that has been built from the past CPT and MASQ studies. The Product Norms Database is a rich knowledge database that helps us to understand product performance over time and how our products are positioned in the market. All new UPT studies will be benchmarked against the normative database as a standard procedure to show the performance of the tested product against the norms.

The UPT will have two main uses within BAT:

1. **Ad-Hoc testing** – any new BAT blends should first undergo the un-branded testing to understand the pure product performance and standing against the norms. All Ad-Hoc UPT tests will contribute to the Product Database. Ad-Hoc testing is mainly used for testing new BAT prototypes against current BAT products and competitor benchmarks. The most common objectives are:
   * ***New Launch****– to find out if the tested prototype achieves required performance against a set benchmark*
   * ***Product Improvement, re-Launch*** *– to find out if test product is superior to the current product on required parameters*
   * ***Product Rationalisation*** *– to ensure that a change in the product does not lead to erosion of product performance*
2. **Product Monitor Programme** –replaces the earlier PSM testing and also introduces a programme where every 3 years (or when necessary if there are significant changes in the market place) we test un-branded a larger set of competitor and BAT products to update norms in our product database.

The Product Monitor is mainly planned for testing existing products in the market and understanding their performance. It is also possible to add some new developmental products to the test when conducting larger landscaping exercise.

With the increased amount of product testing tools and the use of normative database – it is essential to do good planning for the product testing to avoid duplication and to ensure that ad-hoc and monitoring tools are complementing each other.

Any larger landscape exercises – such as previously the PSM – need to be planned in cooperation with the Regional Product Centres and contribute to the agreed Regional Product Strategy.

During 2013 BAT is conducting a large initiative called “G4S – Go for Superiority” and this testing is done un-branded under the umbrella of a Product Monitor.

In addition to the regular maintenance of norms and the recommendation of running the Product Monitor every three years there are also special cases where a market may need to initiate a comprehensive Product Monitor programme.

*For Example: If a market is undergoing a large product development due to legislative changes this would initiate a Product Monitor programme where all the key products need to be re-tested to establish the norms.*

# How to setup un-branded product test?

Un-branded product test (UPT) is a quantitative study where we place each product with the respondents for 4 days home placement. The products are placed in un-branded format – either produced without branding or masked – in white test packaging without any branding information or elements. The principle of the UPT is that the respondent does not know which brand they are smoking and this allows us to get a pure product read.

It is important to note that apart from a couple of new features – the UPT is very similar to the earlier CPT/PSM testing.

If there are products which cannot be masked and the respondents would be therefore recognising the brand they are smoking – UPT method cannot be used as the results would be biased. In such cases the option is to do a Branded Product Test (BPT). For Example this applies to:

* *Products with unique special filter*
* *Products with unique special branded tipping paper*
* *Competitor Menthol or flavoured products where product will be negatively impacted if packs are opened and cigarettes re-packed*
* *Capsule products where masking will not cover special markings on the tipping or the masking will hinder operating the capsule*

## Research Design:

The research design depends on the number and type of products tested. The following principles need to be taken into consideration:

* The standard mandatory placement time for all products is 4 days. This applies both to the Ad-Hoc testing and the Product Monitor. (Please note that this is a change compared to earlier PSM testing which used to have a 2 day placement). It is critical that we follow with all products the 4 days placement in order to have the even quality and consistency
* The standard design is a *Sequential Monadic* placement – however this can only be done with visually identical products. Visually differentiated products need to be separated to *Monadic* matched panels.
* When we use Sequential Monadic placement only 4 products should be tested by one panel in sequential manner – if there are more than 4 products to be tested we have the following options:
  1. Divide the products into matched panels that can do the testing simultaneously. In such case we need to ensure that the design is balanced with an equal amount of products in each panel. Additionally one product needs to be standard across all the panels to help matching the panels on sensorial parameters.

*For example if there are 6 products to test we can set up 2 matched panels where in one we test Common Product + 3 and in the other we test Common Product + 2. Typically the current BAT product in the market would be chosen as the Common Product that is repeated in both panels.*

However, the matched panel approach will increase the sample size and in some market recruiting for certain user groups could be difficult as well as expensive. In such cases, an alternative approach can be considered:

After 16 days (4x4) the respondent is given a 7 day gap to reset their taste pallet and overcome fatigue with product testing. Depending on how many products there are in general to be tested the products need to be divided as equally as possible across the batches.

*For example if there are 6 products to test the recommended design would be 3 products – 7 days gap – 3 products. This would give a balanced testing experience to the consumer. Naturally all products need to be fully rotated within and across the batches.*

The operational feasibility of the option with 7 days gap will need to be reviewed and planned together with the fieldwork agency. This approach whilst potentially having a lower cost will have the following challenges:

* We will need to have minimum 150 people per panel who will test all the products. The longer the taste takes the higher the likelihood of drop-outs. This has to be mitigates with over-sampling.
* If we are confident that we can use the same people for both placements – then we can go ahead with the direct approach of maximum 4 products – 7 days gap – maximum 4 products.
* If there is a risk or a clear concern that we will not be able to have enough respondents who will be doing the entire test – we will need to match the panels on both sides of the 7 day gap and to place a common product across both. Same principle as we did earlier with the matched simultaneous panels. Ideally here we would use as many of the same consumers as possible but this would allow us to recruit more respondents for the second batch testing. Instead of horizontal matched panel approach this would be a vertical matched panel.

*For example if there are 6 products to test with one Common Product in both stages, the design would be 4 products including the Common Product – 7 days gap – 3 products including the Common Product. This would give a balanced testing experience to the consumer. Naturally all products need to be fully rotated within and across the batches.*

* All the packs need to carry an alpha-numeric code and any required local health warnings.

## Design Scenarios:

**Scenario 1: Testing 4 visually identical products**

All 4 products are masked (or if all are BAT products they can be produced un-branded) and placed in one panel for duration of 4 days each.

Depending on the objective of the test there may be one or more panels testing the products such as OWN franchise and SOB.

**Scenario 2: Testing more than 4 visually identical products**

The recommended setup:

Matched panels – number of panels depends on the total number of products tested. For instance if there are 6 products – setup is 2 matched panels, one common product across both, one panel tests 3+Common and the other one 2+Common.

If a matched panel is not a feasible approach from cost or recruitment perspective the other option is:

Using the same panel and the same respondents – over-sample to ensure 150 respondents who will test all the products. Test first 3 products (3x4 days) then give 7 days’ break and then continue with the other 3 products (3x4 days)

If you are not confident that you are able to retain 150 to the end who have tested all 6 products then the option is:

Use from same panel as many as you can + recruit more for the second batch. However here you will have to place one common product on each side of the break. This is due to the fact that some respondents will only test products after the break – therefore we will have to “match” the panels on both sides of the break to ensure the proper comparison.

**Scenario 3: Testing 4 products with visually differentiated products**

Here the design will depend on the type of products tested.

If all the 4 products are visually different from each other – for instance different format or length – then each product needs to be separated to a monadic matched panel. The panels are matched using Regular Brand rating and also demographic and smoker-graphic variables.

If some of the products are visually identical with each other and some are visually differentiated we can optimise the design with following principles:

* All products that are visually identical to be placed in one panel with Sequential Monadic placement
* All products with visual differentiation to be placed in matched panel

*For Example: We are testing 2 KS products and 2 100’s products. In this case we can set up 2 matched panels where in each we test 2 products in Sequential Monadic manner.*

*For Example: We are testing 3 KS products and 1 longer 100’s product. In this case we will also set up 2 matched panels. In one panel we test 3 products in Sequential Monadic manner. In the other panel we place the 100’s product twice. The reason for placing the 100’s product twice is that we need to have a balance of similar amount of products tested across the two panels. It is especially not recommended to have a situation where in one panel a single product is tested. If the respondent knows in one panel that they are testing multiple products it is typical that they will “hold back” the ratings on the first product to allow for something better or worse to come later. If respondent knows they are only testing one product they will have no pressure on “saving” the ratings for later.*

*Special note for ASPAC markets: The ASPAC Region has setup a special pilot programme where in a matched panel setting for visually differentiated testing a common product is placed across all panels 7 days afterwards. This product would be an additional product (can be one of the tested prototypes or a separate product selected by the RPC) and the results are not used for any reporting/analysis. In addition to this the regular brand rating is captured from the memory to aid the panel matching. This approach is used in all the visually differentiated tests within ASPAC until further notice.*

**Scenario 4: Testing Capsule Products**

When testing Capsule products un-branded it is critical first to check if indeed the products can be masked or produced in un-branded format. If sufficient masking is not possible and consumers would recognise the brand – the testing has to be done with the Branded Product Test (BPT).

The second thing that needs to be considered when testing Capsule products un-branded is whether there needs to be separate testing for the “pre-crush” and the “post-crush” product performance.

“Pre-Crush” testing requires a setup where the respondents are not aware that the product contains a capsule. This can only be done if the products can be sufficiently masked or produced without markings that reveal the capsule inside. *(Please remember the first pre-requisite is that consumer does not recognise the brand – this here is a second pre-requisite to make sure the consumer does not recognise the product as a capsule product)*

In “pre-crush” testing it is typical that some consumers will accidentally crush the capsule when smoking the cigarette. We need to be able to identify these consumers and remove them from the sample as the taste of the cigarette would have now changed. This has implications to the questionnaire *(need special questions to check if people recognised a certain taste like menthol)* and also on sample size as we need to over-sample respondents to ensure that in the end we have the sufficient size sample of people who have done a clean test.

“Post-Crush” testing is done by explaining to the respondents that the product contains a capsule and giving sufficient guidance on how to operate the capsule.

If it is necessary to do both “pre-crush” and “post-crush” testing within one project we can use the same respondents for that however it is not possible to rotate the products. By default we have to test the “pre-crush” before we tell the respondents anything about the capsules.

Due to these challenges – when planning any capsule testing please review carefully with the agency the feasibility on masking the branding and capsule marking as well as what are the testing requirements. The final design and complexity depends on these various factors.

**Scenario 5: Testing Menthol Products**

Menthol products (or other distinctively flavoured products) are another type of product where the testing requires special consideration.

It is not possible to test competitor’s menthol products in un-branded format due to the fact that in the masking process the products would lose some of the menthol and this impact the product performance and taste.

It is possible to test BAT menthol product in un-branded format if the products are produced un-branded in the first place. If the objective is to test BAT prototypes against a current BAT product then this method should be used. Also if there are a large number of BAT prototypes to be tested it is recommended to use this method for the short-listing and select the best product(s) to be tested against competitor in another test.

If the requirement is to test BAT products against competitor products the basic testing must be done in Branded format with the BPT. However in order to ensure maximum learning and to get some information on the un-branded pure produce performance against competitors a new approach has been developed with ASPAC Region and this approach is recommended also to be used beyond ASPAC:

This approach combines the use of internal BAT expert panel and consumers with the following 3 components:

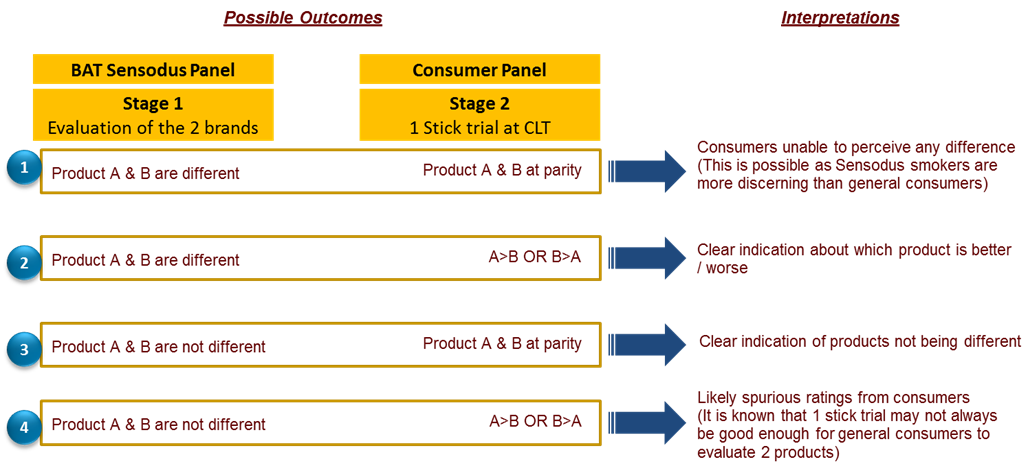


* **BAT Sensodus Panel:** will evaluate the differences and similarities between the test products. Please note that the Sensodus panel tests factual differences in the products – they do not make judgement which product is the best or tastes the best.
* **Consumer Panel – 1 Stick trial at CLT** – This testing is done in blind format with a single masked stick that is masked right before the test at the CLT to avoid any loss of menthol. The pre-requisite here is that it is possible to mask the sticks at the CLT by using a regular masking tape.
* **Consumer Panel – Branded Product Test BPT –** as per the regular BPT guideline – can be done with the same consumers who did the CLT testing. Please note though that the BPT measures a different thing than the UPT so it is not a direct replacement. Please refer to the BPT guideline for the instructions.

Please note that the Sensodus test and the 1-stick CLT test are NOT in isolation pure forms of product testing. They are additional modules that will help us to understand better the performance of the menthol products and the ratings given by the consumers in the BPT test. The 1stick CLT results are not comparable to UPT results and the 1stick CLT test should never be setup in isolation.

How use the non-branded elements:

*For Example: Testing 2 Menthol products (one BAT and one Competitor) – products A and B. The “overall acceptability” rating across the Sensodus panel and the 1-stick CLT trial will help us to confirm how different or similar the products are and if the first impression of the pure product performance.*

****

## Sample size and type:

The basic sample size for a UPT test is 150 smokers per panel. If any sub-group reporting is needed the sample size for each group needs to be 150. A typical test has 150 OWN and 150 SOB smokers to start with.

Sampling should be quota based sampling. It is also possible to use pre-recruited smokers – it can be through formal consumer panels or through re-contact of respondents recruited for previous tests. New guideline for re-contacting respondents is set at a minimum of 3 months after a previous tobacco interview. This has been shortened from the previous guideline of a 6 months gap.

UPT testing will be built on a normative database that will allow us to monitor our products performance over time in the market place and against competitors. In order to do this in a consistent manner the Target Groups that will be used for the testing need to be setup for each market and held as constant as possible over time. Unless the target group definition is kept constant – it will not be possible to do testing against the normative database.

The UPT testing is done separately among OWN and SOB smokers. The exact panel requirements will depend on the business question at hand however here are some typical scenarios:

* For Product Monitor testing (and G4S) the standard approach is to do a separate evaluation by “OWN” and “SOB” smokers. The Product Norms database holds separate norms for the “OWN” and “SOB” ratings.

For all the Strategic Brands (GDB + Key Regional and Local brands) both “OWN” and “SOB” panels need to be included in the test. If there are brands that have a very small base of “OWN” smokers and it is not feasible to recruit a sufficient amount of them – the testing can be considered only among “SOB”. This needs to be evaluated and discussed with key stakeholders case by case.

* For the Ad-Hoc testing the target group depends on the business question and scenario. If the objective is to make a change to a current product in the market it is essential to test the new prototypes with Current “OWN” franchise to find out whether the new prototypes are acceptable alternative for the current franchise. Additional “SOB” panel can be added to the test if at the same time the objective is to find a product which is also an improvement for the potential “SOB”.

If the objective is to launch a new brand in the market naturally the testing would be done among potential “SOB”.

Overall as the addition of panels has a big impact on cost and timings for the project it is important to review in advance which panels are essential for the decision making. Panels and reporting heads that are not relevant to the decision making should be left out.

* 1. **Rules around target group setting:**

**OWN panel:** The target group should cover 80% of the Current Franchise of the particular SKU in order to qualify for a standard OWN panel.

**SOB panel:** The standard SOB panel should be a good representation of the expected SOB in the market. The following things should be considered:

* Clear format rejecters should be left out – not feasible to test the product amongst people who refuse to smoke it. Format acceptance should be tested in the recruiting phase.
* The Target Group for a particular brand should be carefully set up considering current market interactions and strategic objectives. This Target Group definition should be maintained as intact as possible for all subsequent product tests for this particular brand. As things will change in the market place the main thing is to ensure that in relative terms the target group is comparable.
* The Target group should be a mix of relevant SOB brands and not only a single brand *(unless there is a special situation that one large brand represents the entire SOB).*
* The weighting factor for user groups needs to fall within the range of 0.7-1.3 – therefore for small brands, instead of setting quota by individual brand level, sample size can be defined as “group of brands” with similar positioning. The minimum sample size of each such “group of brands” should be at least 25 respondents.
* It is very important that a good balanced and constant SOB definition is established for all of our key brands in each market – please make sure this is carefully reviewed with the agency before testing starts.

*For Example: If we have several small SOB brands in the Full Flavour Premium segment (share of individual brands is less than 1%) users of all key Full Flavour brands can be combined together and given a “sum” quota of 25 smokers. Here we would not apply a quota per brand.*

It is possible to add *single brand target group boosters* to the SOB panel – however the standard SOB panel should always be the starting point and is going to be the requirement for the Product Monitor programme. If there is a business requirement to add a single brand target group that can be arranged as an additional reporting head by boosting the sample.

* 1. **Stimulus materials and requirements:**

The un-branded product test requires a set of non-branded materials:

* *White packs – with test codes*
* *Un-branded or masked sticks*

Amount of stimulus materials needed:

* Each test product will be needed for 4 days placement
* Number of packs placed per test offer would depend on the Maximum Daily Consumption of the respondent
* We should always reserve an extra buffer of 10% for each test product

*For Example: Average Daily Consumption of target group is 19.8 sticks per day. Sample size is 150 smokers. The amount of products needed: 19.8 x 4 x 150 = 11880 sticks – divided by 20 to get 594 packs.*

*Additional buffer of 10% added to this – 594 packs x 1.1 = 654 packs.*

*As we can only include full packs it is good to also make the calculation the other way around to ensure enough products. If in doubt it is better to increase the buffer up to 15% rather than risk running short of products.*

## Product masking:

The cigarette is a product destined for human consumption, requiring precise design and production parameters to guarantee product quality. It is a product which is strongly affected in physical/visual quality as well as smoking characteristics when exposed to temperature or inadequate handling.

Sensory testing requires special controls to avoid bias or reduction of sensitivity. Most of these controls depend directly upon, or are affected by the physical setting in which the tests are conducted.

All products used in UPT studies must be presented fully blind i.e. masked and coded sticks, white and coded packs. The objective is to avoid brand identification that could produce biased results.

Packs should include health warnings in all markets and meet any other legal requirements.

Cigarette masking must be approved standard i.e. non-peelable and fully opaque. The optimum route to achieving this is ECUSTA proprietary masking tape applied on machine via converted HAUNI tip attacher. These guidelines are of course standard good practice to ensure visual equivalence of products in any product testing.

## Product coding:

It is most important that possible biases attributable to different coding effects are minimised. Codes should be constant per product but designed to minimise bias effects. To achieve this, the following rules apply:

* All codes should be alpha-numeric and use one letter and 3 numbers (e.g. **D845**)
* Extreme letters of alphabet and extreme numbers of the 0-9 range should be avoided (e.g. **Z910, A091**). In effect this means avoidance of **A, B, C, X, Y, Z** and **O, 1, 8, 9**.
* Repetition of numbers should not occur (e.g. **X224, T333**).
* Occurrence of odd and even numbers should be balanced as much as possible (e.g. **S346**, **J627**).
* Sum value of each 3 number sequence should be relatively balanced (e.g. **P384**, **F825**).
* Where the first number of the code is greater for one code than the other, this should be offset with the order of occurrence of the letter part of the code in the alphabet (e.g. **R276, D843**). In this case, the further into the alphabet a letter is, the "heavier" it is, thus having the effect of adding more subjective weight to a low code start number.
* Sequences of rounded appearance letters and numbers and sequences of angular appearance letters and numbers should be avoided (e.g. **L471, Q386**).
* Codes with the first letter being equal to the first letter of the brand/product under research should be avoided (e.g. **M672 for Marlboro, D843 for Derby**).
* Sequenced 'runs' of letters and/or numbers should be avoided in multi-product evaluation (e.g. **A234, B456, C678**).

An example of coding for four products to be tested simultaneously is:

|  |
| --- |
| **D843 vs. J672 vs. S347 vs. W275** |

Common sense should apply to coding choices, but if these rules are borne in mind, they will help to minimise any bias effects.

Where supplementary brand identification occurs down the cigarette rod, this can only be obliterated with Chinagraph pencil and whatever masking routine is applied to one cigarette must be reflected for all other product types.

## Product handling

To maintain adequate conservation, the product should maintain a humidity of 60% (+/-2) and a temperature of 22C (+/-1). This requires that the product be stored in air conditioning, in environments that do not receive direct light or severe changes in temperature.

The transport of the product should be carried out in tightly sealed coolers thus avoiding exposure of the product to direct light and/or elevated temperatures for more than 10 minutes. It is preferable to transport the product in air conditioned vehicles.

The product should be transported in small quantities estimating the amount to be delivered daily and quantities should not be loaded above those which would not guarantee proper delivery and handling.

The supervisors and/or interviewers should ideally carry a thermal bag with the daily delivery, thus eliminating exposure of the product to the aforementioned situations.

The good physical quality of the product when received by the consumer has a high impact in its perception. Under no circumstance should the consumer be handed open packs or packs that have been damaged (dirty, torn or wrinkled).

The research/fieldwork agencies should under no circumstances change the codes, the product or the packaging, or have access to the product for private consumption, as well as not alter, change, or modify the delivery methodology of the test.

# Interview and questionnaire flow

UPT is a multi-visit test that is done with pre-recruited consumers. Typically the consumers are first recruited by using various methods and quota sampling. Product Testing Consumer panels are also an option especially in markets that will conduct a lot of product testing.

Interviews are normally conducted at respondents home or chosen location. It is mandatory that the first interview is conducted face to face – however it is possible to administer the follow up interviews on telephone or on-line.

The test products are given to the respondent always one product at a time and any excess products are collected back. If face to face interviews are used, the interviewers are usually delivering and collecting the products. If the follow up interviews are conducted via telephone or on-line the product delivery needs to be arranged with a courier.

The number of visits in a test depends on the number of products tested.

## Visit and Interview flow:

**VISIT 1:**

EXPOSURE OF TEST OFFER 1

The first interview covers orienting the respondent, performing a familiarisation exercise with the ORACLE Sensory Attribute Set and placing the first product. The key requirement is to ensure that the respondent understands the sensory questionnaire and the principles of scaling. The consumers will be given a reference booklet with the test products which explains the sensory attributes and allows them to take notes during the testing.

Additionally, in case of testing visually differentiated products (with matched panel approach without any common product across panels), at the first visit, respondents will be asked to rate their own main brand from memory using the standard SAS list.

**VISIT 2:**

POST SMOKING - RATINGS OF TEST OFFER 1

EXPOSURE OF TEST OFFER 2

The 2ndinterviewshould take place within 4 days' time +/- 8 hours after placement of the 1st test product.

The second interview has the objective of collecting surplus product, collecting back the reference booklet, administering the sensory questionnaire to evaluate the first product placed, de-briefing any specific difficulties and placing the second product for evaluation.

The questionnaire should be completed without smoking a test product simultaneously. This is because the aggregate impression of the 4 days smoking experience is required. The second product is again placed for follow up after 4 days. To ensure consistency of approach, the questionnaire should always be completed by the interviewer -not the respondent himself.

**VISIT 3-5:**

Same process repeats for all test offers up to 4 products tested. A 4-product test would have 5 visits in total. If there are more products to test then either a matched panel approach should be used or after the 4th product there needs to be a 7 days gap before placing further products with the same respondents.

## Questionnaire flow:

**VISIT 1**

**EXPOSURE OF OFFER 1**

|  |
| --- |
| Placing Consumer Reference Booklet |
| Regular brand product rating (for testing visually differentiated products) |

**PLACEMENT FOR 4 DAYS**

**VISIT 2**

|  |
| --- |
| **POST SMOKING – RATINGS OF THE OFFER PLACED** |
| Taste Familiarity |
| Overall Likeability |
| Spontaneous Likes/ Dislikes |
| SAS Attributes |
| Filter& Stick Quality |

**VISIT 3-5 – Similar to Visit 2**

**SAS ATTRIBUTES:**

The basic core questionnaire for Unbranded Product Test remains the same for all different types of products – however the SAS attributes list has been adapted to different product types.

The following SAS attribute lists are available:

* *Conventional Non Menthol product*
* *Conventional Menthol product*
* *Menthol Switch*
* *Menthol Boost*
* *Switch Capsule with other flavours*
* *Boost Capsule with other flavors*

It is very important to specify in the beginning of the project what types of products are tested as this will have an impact on the test design and the questionnaire.

If we are testing visually differentiated products the following needs to be taken into account:

In addition to the SAS+ additional questions will be asked on:

* Relevant image attributes
* Relevant parameters to capture usage / behaviour with the particular product. Here the parameters would be customised as per what is being tested.

## SAS+ attributes – local translation and validation:

The SAS attribute list is the core of the UPT questionnaire. A list of sensory attributes that is kept constant throughout the studies and stored in the normative database.

Whilst the global list has been defined and validated in English – before a market uses the list (or any new attribute) in a study the following two things need to be done:

1. **Translation of the attributes to local language(s).** The most important thing is to capture the meaning rather than try to translate directly from English word by word. Once you have translated into local language is very important to ask a third party to perform a “back-translation” to English to check that your local translation truly is reflective of the English attribute.
2. **Validation with local consumers**. The translated attribute list needs to be tested with consumers to check if they understand the attributes and their meaning in the way that we intended. There has to be a clarity what consumers understand when looking at the attributes. Validation can be done by testing

Please make sure that the SAS+ list is properly translated and validated for your market. Never use an attribute in an UPT unless it is validated and it is clear how consumer understands it.

## Questionnaire length:

The actual questionnaire length varies by market and Target group however we can roughly say that:

Visit 1 = 5-10 minutes

Visit 2 & Subsequent Visits = 15 minutes for testing conventional products; 15-20 minutes for testing Switch/ Boost products.

For face to face interviews some additional time has to be accounted for the basic meeting and greeting.

# Key Metrics and Analytics

## Key outcome measures:

There are 4 main areas of analysis:

* *Product Performance*
* *Consumer Pathway to drive overall acceptability*
* *Product Attribute Space Map*
* *Product Optimization and Penalty analysis*
* *Comparison against normative database*

### Product Performance:

This section summarises the product scores and gives a snapshot of the ratings in the study. Ratings are done separately among OWN and SOB smokers and the main outcomes are:

* *Overall Acceptability ratings*
* *Magnitude and “Just Right” summary*
* *An overview of the magnitude and “Just Right” ratings signifying superiority/ inferiority by marking the significant differences for the attributes.*
* *Detailed findings of the ratings*

Results should be displayed as product profiles supported by conventional mean, top box/top 2 boxes and standard deviation with splits per attribute, product and consumer segment, with tests for significance of difference.

All product tests must include Acceptability rating. Tables with means, top box/top 2 box and standard deviations for each attribute should be produced by smoker group indicating significant differences: NS, 90% (\*), 95% (\*\*), 99% (\*\*\*).

Standard confidence level that BAT uses with 150 sample size is 95%.

Additionally, it should be indicated in the table which means are significantly different and which are not by using letter suffixes as shown below. Means coded with the same letter indicate that there is no statistical difference between them:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **PRODUCT** | **1** | **2** | **3** | **4** | **5** |
| **MEAN** | **2.75a** | **2.48b** | **2.69a** | **3.01c** | **2.59ab** |

*In the above example of comparison among five products, there is no significant difference between products 2 & 5; no difference between products 1, 3 and 5, but products 1 & 3 are significantly different from product 2; and product 4 is significantly different from all other products.*

For 'Just Right Scales', the proportion of respondents rating the product at ‘3’ should be tested for significant difference between the products. Significance should be indicated by using \*/\*\*/\*\*\* as superscript where appropriate.





### Consumer pathway to drive overall acceptability:

The consumer pathway is derived by using Structural Equation Modeling (SEM). This analysis is used in the Product Monitor as a standard deliverable. It can also be used in Ad-Hoc studies where the target group is set up with the broader definition (same as the Product Monitor). The analysis helps us to understand which attributes are driving the overall acceptability of the product – what is important for the consumer and what is less important.



### Product attribute space map:

The product attribute space map is plotted with a correspondence map. This can be done for tests that have more than two products. The analysis is done separately for “OWN” and “SOB”. The correspondence map will help us to understand:

* How the products are associated with the attributes?
* How the products are differentiated from each other?
* How the attributes associate to each other?



### Penalty analysis:

Product optimization consists of input on how to improve the product or how to make it more differentiated. Penalty analysis would tell us the risk associated with moving away from ‘Just right’. It would suggest us the most sensitive attributes, the penalties associated with it and the direction to move (to increase or decrease the extent).The analysis is done separately for “OWN” and “SOB” for each product at a time. Below is an illustrative example of Penalty analysis, which indicates that the strength of taste is the most sensitive attribute.

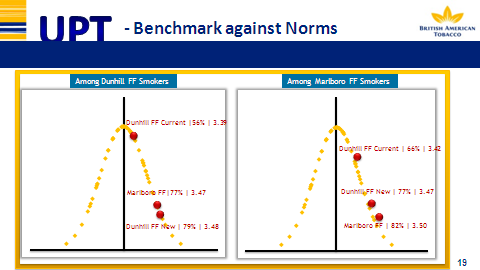


### Comparison against Normative Database:

In addition to evaluating test products against a pre-decided action standard, performance will be benchmarked against the Product Normative Database. This will help to understand how the tested products are performing in the market context against the norms. *While we already have the capabilities in to benchmark against the Product Normative Database, please note that the availability and type of norms for your particular market needs to be checked in advance when planning the study. Availability of norms depends on previous tests done in your market and that they have been loaded into the database. In case not sufficient market specific norms are available for your project the research agency will review the availability of norms from the same Region or even Globally and propose what can be used instead.*

How do we compare against the norms? The key output from the Product Normative database comparison is to specify the position of test products in terms of a “percentile” - which indicates where a specific product stands in comparison with the best performing (99.99th percentile) and the worst performing (0.0th percentile) product tested in that market.

*For example in the chart below we can see that among Dunhill FF smokers, with a mean score of 3.39, the product stands at 56th percentile (in other words: of all products tested in this market, this specific test product stands within top 56%). Similarly, among Dunhill FF smokers, Marlboro FF product stands at 77th percentile and Dunhill FF New prototype stands at 79th percentile.*



### BAT Product Normative Database (C-PSI):

C-Psi stands for “Consumer – Product Survey Interface”. This is an interactive database that has been designed to capture and store data from the BAT product surveys. The database holds currently the results from our 2006-2012 CPT and MASQ studies. Please note that only studies that had a 4 day placement have been included.

The database allows for the calculation of norms and to extract data for advanced statistical analysis.

IMRB International is the agency in charge of the development and management of the C-Psi database in close cooperation with BAT Global Product and SP&I teams. Going forward benchmarking is a standard deliverable in a UPT study. If the study is not coordinated by IMRB International – the coordinating agency needs to liaise with IMRB to deliver them the test data and to obtain the norms.

It is also possible for relevant BAT employees (Product or SP&I) to get access to the C-Psi database. Please contact Global SP&I team for more information on this.

## ACTION STANDARD SETTING:

The action standards for the UPT depend on the business objectives of the test – typically the objective is to achieve superiority or parity against a given benchmark.

For Ad-Hoc testing the action standard setting depends on whether the testing is done only among OWN or SOB or both.

For Product Monitor the standard objective is to achieve superiority against the defined competitor benchmark(s).

*The definition of product superiority is based on the relative performance of our products against competitive products, among “OWN” and “SOB” smokers.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **SOURCE OF BUSINESS SMOKERS** | | |  |
|  | BAT PRODUCT  INFERIOR TO | | BAT PRODUCT  PARITY WITH | BAT PRODUCT  SUPERIOR TO | |
| **OWN SMOKERS** | BENCHMARK | | BENCHMARK | BENCHMARK | |
| BAT (OWN) PRODUCT INFERIOR  TO BENCHMARK | **UNACCEPTABLE** | | **UNACCEPTABLE** | **UNACCEPTABLE** | |
| BAT (OWN) PRODUCT PARITY  WITH BENCHMARK | **UNACCEPTABLE** | | **PARITY** | **SUPERIORITY** | |
| BAT (OWN) PRODUCT SUPERIOR  TO BENCHMARK | **PARITY** | | **SUPERIORITY** | **SUPERIORITY** | |

While the basic construct of defining the performance of BAT products versus competition has remained same as earlier, the following changes need to be noted:

Testing will be done among OWN smokers of BAT product and broad-based SOB target group (instead of having only one brand smokers in the SOB TG). This means that the outcome has to be interpreted as follows:

* *Performance of BAT product versus the benchmark product among BAT OWN smokers (same as earlier approach)*
* *Performance of BAT offer versus the benchmark product among relevant broad SOB (different from earlier approach where the benchmark product smokers were used as a TG)*

This construct will allow setting action standard against multiple competition products as long as they have been tested among broad-based SOB smokers. If a product test is conducted among one specific SOB brand users the construct will remain the same, but the outcome will indicate the relative standing of two products head-to-head, instead of reflecting the performance among broad-based SOB.

In case one specific SOB brand gets over-represented (more than 30% of total SOB), additional sample size is needed to conduct the analysis among non-regular users of that particular SOB brand. This is due to the fact that with such high proportion of regular smokers included in the panel construct, the outcome is likely to get influenced in favour of the particular brand.

Testing among broad based SOB is overall preferable as it gives a more holistic understanding on the acceptance of the product in the market.

For Ad-Hoc testing the standard objective is to achieve superiority with newly produced prototypes against the current product.

The action standard would be to achieve superiority (or parity in some cases) with the new prototypes versus existing product among OWN and/or SOB smokers.

If the objective is to change the existing product in the market the acceptance of OWN smokers is a must and we commonly also add the SOB panel. However there may be cases where the SKU in question is very small so it is not feasible to find 150 OWN smokers. In such exceptions it may be acceptable to do the testing only among SOB smokers.

# Agency for Un-branded Product Test

IMRB International is the lead agency responsible for managing the BAT product testing and the Product Normative Database (C-Psi).

Any **Product Monitor** Testing needs to be done with IMRB International.

Any **Capsule testing or visually differentiated product testing** needs to be done by IMRB International.

**Any testing that requires the SEM (Structural Equation Modelling) analysis** **needs to be done by IMRB International.** SEM is included in the Product Monitor tests. If SEM is a requirement in Ad-Hoc testing then the SEM analysis has to be done by IMRB.

Standard Ad-Hoc UPT excluding SEM can be done with a local agency. If a local agency is used they need to liaise with IMRB to get the benchmarking done against the Product Normative Database. In order to do this the study results need to be delivered to IMRB International for inclusion in the database and to receive the norms.