E

TWO/42/3 Add. Rev.
ORIGINAL: English
DATE: September 9, 2009

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY FOR ORNAMENTAL PLANTS AND FOREST TREES

Forty-Second Session
Angers, France, September 14 to 18, 2009

REVISED ADDENDUM TO TGP DOCUMENTS

Document prepared by the Office of the Union

1. The annexes to this document provide the following information concerning TGP documents:

Annex I: comments provided by an expert from Denmark concerning document TGP/8/1 Draft 13; and.

Annex II: comments on the TGP documents made by the Technical Working Party for Vegetables (TWV), the TWC and the Technical Working Party for Agricultural Crops (TWA).

2. Annex II of this document replaces document TWO/42/3, Annex I.

[Annexes follow]

ANNEX I

COMMENTS ON SECTIONS 5 AND 6, PART II OF TGP/8/1 DRAFT 13
SENT BY MR. KRISTIAN KRISTENSEN

Section 5.2.2.: I suggest that "relevant criteria" to be replaced with "statistical assumptions".

Comment [c4] in section 5.3.1. I am not sure how this can be moved to Part I because this note does not apply to all statistical methods but only to those were the random variation is determined by the underlying distribution (usually pure Poisson, binomial or multinomial), i.e. here to the tests in contingency tables (either based on Chi-square tests or Fisher exact test). So this note does not apply to analysis of variance methods or use of LSD-values when data from individual replicates or years are analysed assuming data to be normally distributed.

Section 5.3.8. Here it is stated that: "Hence the distribution of scores in different classes observed for this reference variety is considered to be the expected distribution". I do not think that the values in this column can be called expected values as they are samples from a large population and may thus deviate from the expected distribution. So I think that this section should be reformulated.

Section 5.3.9. As a consequence of my comments to section 5.3.8 the value in the table and the calculations are wrong. I suggest deleting the whole section.

Section 5.3.11. I suggest the text to read "For comparing reference variety 1 with the first generation of the candidate we get:" Alternatively the aggregated value of both candidate varieties could be compared with the reference variety.

Section 5.3.11. As a consequence of my comments to section 5.3.8 the formula for the Chi-square should read: (34-23.51)^2/23.51+(6-14.82)^2/14.82+(6-7.67)^2/7.67
+(12-22.49)^2/22.49+(23-14.18)^2/14.18+(9-7.33)^2/7.33=21.05
(using the formulas in section 2 of TWC/27/14).

Section 5.3.12. If kept I suggest writing: "When comparing two varieties the number of degrees of freedom are one less the number of classes."

Section 5.3.14. Using the method given above (and described in TWC/27/14) I do not get the revised values.

Section 6.1.6 to section 6.1.11. I have just realised that the description of the test seem to be in error. The reason is that the description seems to focus on the probability of getting the actual outcome if the distribution is the same for both varieties (section 6.1.6). I think that the probability should be calculated as the probability of getting a result that is at least as extreme as the actual one (which is similar to the probability that is calculated when using e.g. t-tests for normally distributed data). For the example in table 1 I think that there are in total 4 outcomes that are at least as extreme as the actual one - and have the same marginal’s. These are the following:

4 9
8 3,

3 10
9 2,

2 11
10 1 and,

1 12
11 0

The first one is the actual one. The probabilities of each of those 4 outcomes are 0.04362, 0.00582, 0.00032 and 0.0000048. Summing those 4 values I get 0.04977 which rounded to 2 decimal points are 0.05 and thus different from the result stated in section 6.1.8. Note that I have here assumed a one-tailed test. (If a two tailed test is required there are more outcomes that are as extreme as the actual one and the p-value gets larger.)
For the example in section 6.1.10 I think that there are 2 outcomes that there are as extreme as the actual one:

1 9
11 3 and

0 10
12 2

These have probabilities 0.001346 and 0.000034 which sums to 0.001380.
This viewpoint is supported by several programs that can be found on the internet, e.g. *http://faculty.vassar.edu/lowry/fisher.html*.

Section 6.1.10: The fraction line is missing in the expression for p.

[Annex II follows]

ANNEX II

COMMENTS ON TGP DOCUMENTS MADE BY THE TECHNICAL WORKING PARTY FOR VEGETABLES, THE TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS AND THE TECHNICAL WORKING PARTY FOR AGRICULTURAL CROPS

The following comments on draft TGP documents were made by the Technical Working Party for Vegetables (TWV), at its forty-third session, held in Beijing, China, from April 20 to 24, 2009, and by the Technical Working Party on Automation and Computer Programs (TWC), at its twenty-seventh session, held in Alexandria, Virginia, United States of America, from June 16 to 19, 2009.

(a) New TGP documents

TGP/8 "Trial Design and Techniques Used in the Examination of Distinctness, Uniformity and Stability"

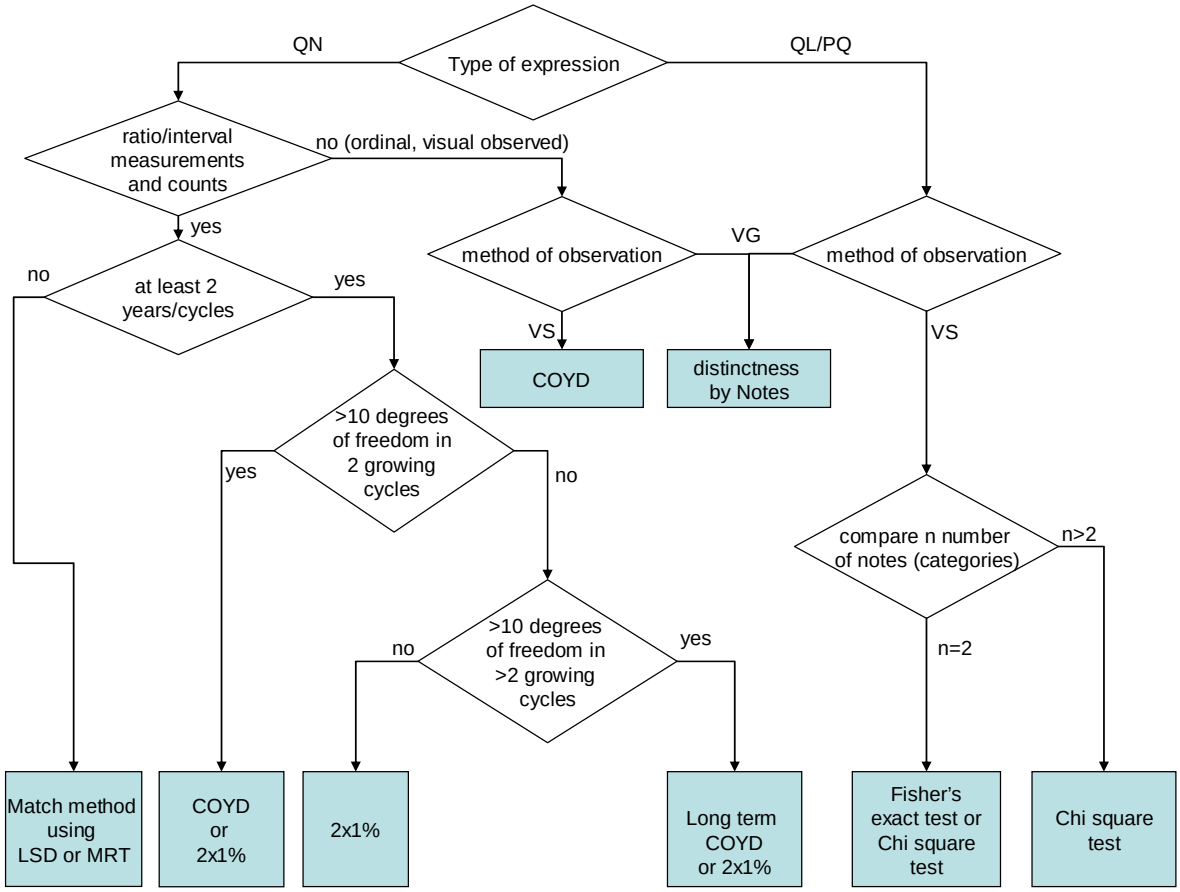
Comments on document TGP/8/1 Draft 13:

Introduction	to read "PART II: TECHNIQUES USED IN DUS EXAMINATION: includes, in particular, details on certain techniques referred to in documents TGP/9 "Examining Distinctness", and TGP/10 "Examining Uniformity".	TWC
Introduction	To read: "PART II: TECHNIQUES USED IN DUS EXAMINATION: provides details on certain techniques referred to in documents TGP/9 "Examining Distinctness", and TGP/10 Examining Uniformity where further guidance is considered appropriate. ~~It should be noted that the techniques included in Part II are not the only techniques that are suitable for use in the DUS examination. For example, DUS expert observation is an important technique but is not included in document TGP/8.~~"	TWA
General	to provide an explanation of the term "reference variety" throughout the document (e.g. COY)	TWC
	PART I	
1.1	text in square brackets to be deleted	TWC
1.3.1.1	Last sentence to read: "…However, for example, it may be considered appropriate to conduct tests at more than one place for the following purposes:" To explain the need that before using more than one location the interaction genotype-environment has to be considered	TWA
1.3.1.2	to delete "s 1.2.2.5 and"	TWC
1.3.2.2	to read "If multiple growing trials are used as explained in sections 1.3.1.(a) and (c), DUS could be examined at all growing trial locations. However, in general, DUS is not examined at all growing trials locations."	TWC
1.3.2.2	To be deleted	TWA

1.3.2.3 Title (a)	To read: "(b) Additional tests ~~DUS examined using characteristics examined at different locations~~"	TWA
1.3.2.3	To read: "For example, additional tests (see section 1.6) *[cross ref.]* may be carried out to examine particular characteristics e.g. greenhouse tests for disease resistance, laboratory tests for chemical constituents etc. In such cases, the data for particular characteristics can be obtained at a different location to the main growing trial. ~~In other cases, reserve trial data may be available for some or all characteristics which could not be observed in the growing trial at the primary location. In cases where the data for the characteristic(s) are obtained exclusively from the reserve trial, the situation is similar to that for an additional test, although it would be important to record that the variety description for the characteristics concerned was not based on the normal (primary) location. The situation where data from different locations (i.e. the primary location and reserve location) for the same characteristic are combined is covered in paragraph (c).~~	TWA
1.3.2.4, 1.3.2.5	To delete reference to descriptions in 1.3.2.4 and 1.3.2.5 and make a new paragraph dealing with descriptions	TWA
1.5.2	Title to read: "1.5.2 Number of Plants in the trial"	TWA
1.5.2.1	1.5.2.1 The number of plants~~/parts of plants to~~ in the trial ~~examined~~ is influenced by several factors such as genetic structure of the variety, way of reproduction of the species, the agronomic features and the "feasibility" of the trial. The most significant criteria to determine the number of plants are, the variability within and between varieties, and the method of assessment of distinctness and uniformity.	TWA
1.5.2.2	to be deleted	TWA
1.5.2.3	to be deleted	TWA
1.5.2.4	to be deleted	TWA
1.5.3	To include incompletely randomized trials (to cover grouping) in a future version of TGP/8.	TWA
1.5.3.1.7 (table)	title of third row to read "Variety mean / Statistical analysis of records for a group of plants / [Replicate plots for group data records] / (MG/MS)	TWC, TWA
1.5.3.1.7 (table)	to explain the terms MG, MS, VG, VS	TWC, TWA
1.5.3.3.2	to be deleted	TWC, TWA
1.5.3.3.4.6	second sentence to read "The blocks should be formed so that the variation between plots within each block is minimized.	TWC, TWA
1.5.3.3.7.2.6 (table)	to delete " "	TWC
1.5.3.3.7.4	to be deleted	TWC, TWA
1.6	To be moved earlier in the structure	TWA

1.7	To read: "1.7 Changing Methods Changes in the methods of assessing DUS may have a significant impact on decisions. Therefore, due consideration should be given to seeking to ensure that there is consistency in decisions ~~and that applicants are aware of the changes to the method~~	TWA
2.1.1	Correct paragraph numbering	TWA
2.3	first paragraph to be deleted	TWC, TWA
2.3.1 (title)	to delete "[/variety means]"	TWC, TWA
2.3.1	to delete text in square brackets	TWC
2.3.1.1.3	to add blank line before 2.3.1.2	TWC
3.1	to delete note in box	TWC, TWA
3.2.1.3 (b)	to read "The 2x1% method to assess distinctness, which has also been developed by UPOV to analyze data from two or more years of growing trials where there are at least a certain minimum number of varieties in trial. Differences are assessed in each year using a statistical test based on a two-tailed LSD to compare the within-year variety means. Whether differences are sufficiently consistent is determined by the requirement that two varieties are significantly different in the same direction at the 1% level in both years, or, where trials are conducted in three years, in at least two out of three years. Details of the 2x1% method are given in document TGP/8 Part II section 4."	TWC
3.2.1.3 (b)	Accept the text proposed by the TWC subject to the deletion of the phrase "where there are at least a certain minimum number of varieties in trial."	TWA
3.2.1.3 (c)	to read "The Match method to assess distinctness was developed for use where the trials are conducted by the breeder in the first year and examined by the testing authority in the second year (see document TGP/6 section 2/1). They typically involve relatively small scale trials. Whether differences are sufficiently consistent is tested using a statistical test (eg LSD, MRT, Chi-Square or Fisher's Exact) to gauge whether the differences in the second year are significant and agree with the "direction of the differences" declared by the breeder in the first year. The choice of statistical test depends on the type of expression of the characteristic concerned. Details of the Match method are given in document TGP/8 Part II, Section 5." *Subsequent to the TWC session, the following text was proposed by an expert from Australia in conjunction with the re-drafting of text of Part II, Sections 5.3 and 6 (see comments in table)* to read "The Match method to assess distinctness was developed to analyze data from more than one year of testing . Trials are conducted by the breeder in the first year and examined by the testing authority in the second year (see document TGP/6 "Arrangements for DUS	TWC

	Testing", Section 2 "Examples of Arrangements for DUS Testing"). Whether differences are sufficiently consistent is assessed using a statistical test (e.g. LSD, Multiple Range Test (MRT), Chi-Square or Fisher's Exact) to gauge whether the differences in the second year are significant and agree with the "direction of the differences" declared by the breeders in the first year. The choice of statistical test depends on the type of expression of the characteristic concerned. Details of the Match method are given in document TGP/8 Part II, Section 5" ~~The above methods use different statistical tests to assess whether differences between variety means are significant. The choice of the statistical test that is used has implications for the risks to the breeder and the tester of making statistical errors and is discussed below.~~	
3.2.1.3 (c)	Accept the text proposed by Nik Hulse (Australia) to refer to "growing cycle" instead of "year of testing" or "year"	TWA
3.2.1.3	to delete the words "and is discussed below"	TWC
3.2.1.4, 3.2.1.5	to be deleted and to be replaced by an explanation that different statistical methods will produce different results and to consider that in the context of harmonization	TWV
3.2.1.4, 3.2.1.5	to be deleted and to be replaced by an explanation that "In the context of consistency and harmonization, it should be noted that different statistical methods will produce different results."	TWC, TWA
3.2.1.5	second sentence to be deleted	TWC
3.3 (title)	to read "Summary of selected statistical methods for examining distinctness"	TWV, TWC, TWA
3.3	- title of flow diagram to be amended and to avoid an indication that there is a preference of COYD over 2x1% method if there are more than 20 degrees of freedom - to clarify that other statistical methods would not be excluded	TWV
3.3.1 (table)	- to update the minimum degrees of freedom according to changes agreed for the relevant methods (see below) - to delete "Distribution" column - to replace Chi square and Fisher's exact test with row for Match method - to add a column to indicate method of observation as "MS/VS" for COYD and Long Term COYD, with a note that those methods might also be applicable for MG and VG in certain circumstances; and to indicate "MS" for 2x1% method and "VS" for Match method	TWC
3.3.1	To delete the table and to read "<u>Selected</u> techniques used in DUS examination"	TWA
3.3.1	flow diagram to read as follows:	TWC



PART II		
Title	to read "Selected techniques used in DUS examination"	TWV, TWC
General	to check that the term "clearly distinct" is replaced by "clearly distinguishable", "distinct" or another suitable term (e.g. 6.1.9)	TWV, TWC
1	Second sentence of first paragraph: to check whether the term "originality" is correctly used or should be replaces.	TWA
1.1.3	To read: "A DUS examiner may have a situation where two varieties receive ~~a~~ different note<u>s</u> ~~(e.g. Variety A is Note 3 for a given characteristic and Variety B is Note 4)~~, but the two varieties are considered by the examiner to be similar. The difference could be due to the fact that the varieties were not grown very close each other (i.e. had different environmental conditions), or to variability of the observer when assessing the notes, etc.	TWA
1.3.1	To add the following text at the beginning of the section: "It is important to take care of the correlation between characteristics when weighting. If two characteristics (e.g. two plant heights) are linked, it is advised to use only one of them in GAIA to avoid double weight."	TWA
1.3.4.1	to refer to test guidelines rather than crop guidelines	TWA

3.1	to read "– there should be at least 10, and preferably at least 20, degrees of freedom for the varieties-by-years mean square in the COYD analysis of variance, or if there are not, then Long-Term COYD can be used (see 3.6.2 below);"	TWC
3.1	To maintain the recommendation of 20 degrees of freedom and to include the recommendation of the TWC in a future revision of TGP/8	TWA
3.7	to read "The COYD method can be applied using TVRP module of the DUST package for the statistical analysis of DUS data, which is available from Dr. Sally Watson (Email: info@afbini.gov.uk) or from *http://www.afbini.gov.uk/dustnt.htm.* Sample outputs are given in Part II section 3.10 *[cross ref.]*."	TWC
3.9.2.1, 3.9.2.2	to replace "SE" with "standard error" (3 occurrences)	TWC
3.9.2.5	formula to be centrally aligned	TWC
4. (title)	to read "2X1% METHOD"	TWC
4.1.1	to add indent to read "– there are at least 10, and preferably at least 20, degrees of freedom"	TWC
4.1.1	the TWA did not agree with the TWC proposal to add indent to read "– there are at least 10, and preferably at least 20, degrees of freedom"	TWA
4.2 (title)	to read "The 2X1% method"	TWC
4.2.1	second sentence to read "The tests in each year are based on Student's two-tailed t-test of the differences between variety means with standard errors estimated using the residual mean square from the analysis of the variety x replicate plot means."	TWC
4.2.1	To clarify that it is not the residual of the individual plants what should be used	TWA
4.2.2	to delete final sentence of second indent	TWC, TWA
5.1.4	to delete "or establish that the type of data collected does not fit the parametric assumptions"	TWV
5.1	to be deleted	TWC, TWA
5.2	to be deleted	TWC, TWA

5.3 general	section to be edited according to the comments below, the proposals agreed by the TWC at its twenty-sixth session (see document TWC/26/29 "Report", paragraph 29: items 21, 23, 21) and any written comments provided to Mr. Nik Hulse by July 3, 2009. Mr. Hulse to prepare a new draft of the section for circulation by the Office by July 17, 2009 to the TWC, with a request for comments to be provided July 31, 2009. On the basis of comments received, Mr. Hulse to prepare a text by August 3, 2009, to be presented to the Technical Working Party for Agricultural Crops (TWA) and subsequent Technical Working Party sessions in 2009. *The proposed new draft of Section 5, on the basis above, is presented in the Appendix to Annex I to this document*	TWC
	NEW SECTION (MATCH METHOD) (TWO/42/3 / TWF/40/3, Appendix to Annex I)	
General	Sections 5.1 and 5.2 to be moved as new sections after Section 6 under the title Match Approach	TWA
General	to read "5. MATCH METHOD 5.1 Requirements for application of method 5.1.1 The Match method is appropriate for assessing distinctness of varieties where: - observations made on a plant (or plot) in the second year are compared to observations made by the breeder in the first year. - there are claimed differences between plants (or plots) of a variety based on information from the first year trial - the requirements of the method depend on the particular statistical test that is used (e.g. LSD, Multiple Range Tests (MRT), Chi-Square or Fisher's Exact). 5.2 Match Method 5.2.1 The Match method to assess distinctness was developed for use where the trials are conducted by the breeder in the first year and examined by the testing authority in the second year (see document TGP/6 section 2/1). Whether differences are sufficiently consistent is tested using a statistical test (eg LSD, MRT, Chi-Square or Fisher's Exact) to gauge whether the differences in the second year are significant and agree with the "direction of the differences" declared by the breeders in the first year. The choice of statistical test depends on the type of expression of the characteristic concerned. For two varieties to be distinct using the Match method, the varieties need to be significantly different in the same direction claimed by the breeder in the first year. 5.2.2 The requirements of the method depend on the particular statistical test that is used (e.g. LSD, MRT, Chi-Square or Fisher's Exact). For quantitative characteristics, the statistical test may be based on a one-tailed LSD, if there is one candidate, or on a one-tailed MRT, if there is more than one candidate included in the growing trial. A Chi-square test or Fisher's	TWC

	exact test may be used for pseudo-qualitative or qualitative characteristics where the requirements for those tests are met. 5.2.3. The Match method typically involves relatively small scale trials. The number of candidate and reference varieties in the trial is limited to the most similar varieties of common knowledge. Although these tests are most useful in trials of cross-pollinated varieties, they can be similarly applied to trials of self-pollinated and vegetatively propagated varieties provided the relevant criteria are met.	
5.3 title	to read "Chi-square test applied to contingency tables"	TWC
5.3	To become section 5 Pearson's chi-square test	TWA
5.3	(a) to provide list of requirements and circumstances for the use of Chi-square test applied to contingency tables, which would include: - the only source of variation should be caused by random sampling, e.g. there should be no variation due to soil conditions, etc. - useful where observations on a characteristic are allocated to two or more categories (classes) - the minimum expected value in each category should be five (b) to explain contingency tables	TWC
5.3.2	to read "In some cases, distinctness may be established by classifying individual varieties into broad groups and demonstrating statistically different grouping patterns for different varieties. Such examples include counts based on the flower color groups - red, pink or white etc. and the disease/pest/nematode infection classes. Data based on counts of individuals in a sample/population belonging to each of several classes require a different kind of statistical analysis. A method commonly used for analyzing such enumeration data is called the *Chi-square* (χ^2)."	TWC
5.3.6	to indicate "contingency table" in the title	TWC
5.3.16 to 5.3.19	to be deleted	TWC
6.	section to be edited according to the comments below and any written comments provided to Mr. Nik Hulse by July 3, 2009. Mr. Hulse to prepare a new draft of the section for circulation by the Office by July 17, 2009, with a request for comments to be provided by July 31, 2009. On the basis of comments received, Mr. Hulse to prepare a text by August 3, 2009, to be presented to the Technical Working Party for Agricultural Crops (TWA) and subsequent Technical Working Party sessions in 2009. (as for Section 5). *The following amendments were proposed on the above basis:* **6. FISHER'S EXACT TEST** Fisher's Exact Test is a statistical test used in the analysis of categorical (qualitative) data where the number of samples (i.e. sample size) is small and is named after its inventor, R.A. Fisher. <u>Fisher's Exact test applied to 2 x 2 contingency tables is useful where;</u> <u>- observations on a characteristic are allocated to two or more</u>	TWC

	categories (classes) - the only source of variation should be caused by random sampling, e.g. there should be no variation due to soil conditions, etc. - the expected values in each category are less than 10 **6.1 Assessment of Distinctness** 6.1.1 Fisher's Exact Test is used to determine if there are non-random associations between two categorical variables in a 2 x 2 contingency table and can be used when the sample number for one or more categories for each variety is less than 10 (see bold framed cells in Table 1) or when the table is very unbalanced. Where there is a larger number of samples (i.e. 10 or more), a chi-square test is often preferred. ~~– as it is usually quicker to calculate.~~ […] 6.1.9 Interpreting the p value calculated by Fisher's Exact Test is straight forward. In the example above, p = 0.04 meaning that there is a 4% chance that, given the sample size and distribution in Table 1, observed differences are due to sampling alone. Given the small sample size, and the need for varieties to be clearly ~~distinct~~ distinguishable from each other, it is open to examination authorities to choose p = 0.01 as the upper cut off significance acceptability level of our null hypothesis.	
6.	to provide list of requirements and circumstances for the use of the method	TWC
6	Introductory paragraph to read: Fisher's Exact Test is a statistical test used in the analysis of categorical (qualitative) data where the number of samples (i.e. sample size) is small ~~and is named after its inventor, R.A. Fisher~~. Fisher's Exact test applied to 2 x 2 contingency tables is useful where: observations on a characteristic are allocated to two or more categories (classes) - the only source of variation should be caused by random sampling, e.g. there should be no variation due to soil conditions, etc. - the expected values in each category are less than 10	TWA
6.1.1	To delete: "as it is usually quicker to calculate"	TWA
6.1.2. Example 1	To make it a general example, i.e. not to refer to lucerne	TWA
6.1.4	The TWA did not agree with the proposed deletion in this paragraph in reply to the comments made by Mr. Kristian Kristensen (Denmark)	TWA
6.1.5	To add the following text at the end of the paragraph: "In this case, the probability is calculated as the sum of the probabilities for each possible event that is as larger or larger than the observed. Consequently, in addition to the observed, the number of dark blue flowers that would give a successful outcome would be 9,10 or 11 for Variety 1 and 2, 1 or 0 for Variety 2."	TWA

6.1.9	To have p=0.05 in the second sentence and to replace "distinct" b~~y~~ "distinguishable" in the third sentence.	TWA
6.2	to be deleted and to be covered by new section in future revision of TGP/8 (see below)	TWV
6.2	to be deleted	TWC, TWA
7.1	note in square brackets to be deleted	TWC
7.1.5	Correct paragraph number in the title	TWA
7.1.5.2	To change the population standard to 10% and the acceptance probability to 95%.	TWA
7.1.5.3	second line, to read figures "(1 to 7)"	TWA
7.1.5.4	note in square brackets to be deleted	TWC
7.1.6	Title to read: "Method for one single test"	TWA
7.1.7	Title to read: "Method for more than one single test (year)"	TWA
7.1.8	To delete title	TWA
8.1	to delete "COYU is an appropriate method for use in assessing the uniformity of varieties"	TWC
8.1	To explain the notion of "reference variety" in COY" in the document.	TWA
8.9	to read "The COYU criterion can be applied using COYU module of the DUST software package for the statistical analysis of DUS data. This is available from Dr. Sally Watson (Email: info@afbini.gov.uk) or from *http://www.afbini.gov.uk/dustnt.htm*."	TWC
8.11	to delete paragraph after Table 1	TWC
9.1 Title	to read "Uniformity assessment on the basis of relative variance method"	TWC, TWA
9.1 Introduction	to add an introduction based on Section 9.4.4 and to add a requirement that there should be a normal distribution for the method to be used. *The following amendments were proposed on the above basis:* 9.1 ~~Use of the relative variance method~~ Uniformity assessment on the basis of the relative variance method ~~In Australia, the relative variance method is applied to any measured characteristic that is a continuous variable, irrespective of the method of propagation of the variety.~~ The relative variance for a particular characteristic refers to the variance of the candidate divided by the average of the variance of the reference varieties (i.e. Relative variance = variance of the candidate/average variance of the reference varieties). The data should be normally distributed. The relative variance method may be applied to any measured characteristic that is a continuous variable, irrespective of the method of propagation of the variety.	TWC

9.1	Introduction to read: " The relative variance for a particular characteristic refers to the variance of the candidate divided by the average of the variance of the reference varieties (i.e. Relative variance = variance of the candidate/average variance of the reference varieties). The data should be normally distributed. The relative variance method may be applied to any measured characteristic that is a continuous variable, irrespective of the method of propagation of the variety."	TWA
9.1	to read "The relative variance method may be applied to any measured characteristic that is a continuous variable, irrespective of the method of propagation of the variety."	TWV, TWC
9.1.1	to add space before "∞"	TWC
9.1.2	To be deleted	TWA
9.2.1	Table 1; to delete the rows for sample size 10, 15, 20, 25	TWA
9.3.1	To include guidance on the minimum number of reference varieties to be included in the trial	TWA
9.3.2	To be deleted	TWA
9.4.5	To delete example 2	TWA
9.5.2	Table 4; to delete the rows for sample size 10, 15, 20, 25	TWA
9.6	to be deleted	TWC, TWA

TGP/11 "Examination of Stability"

Comments on document TGP/11/1 Draft 5:

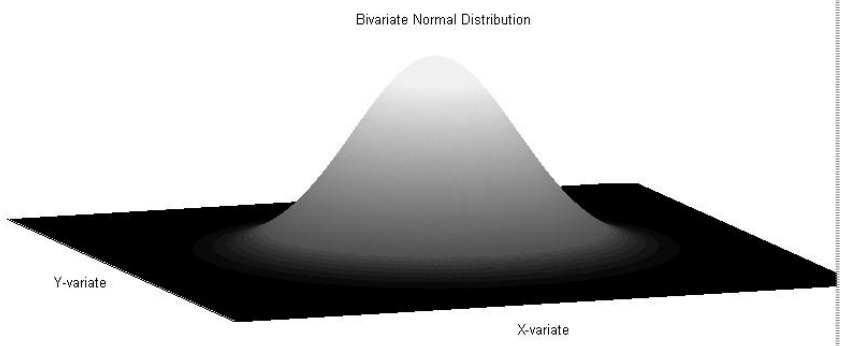
General	The expert from Australia explained that, in Australia, stability was examined for seed-propagated varieties by growing two generations and verifying that there was no difference in the characteristics observed.	TWC
	The TWA noted that document TWA/38/3, paragraph 18(f) explained that, in addition to guidance on the examination of stability through the examination of uniformity, the next draft of document TGP/11/1 should provide guidance on the direct examination of stability, with the assistance of experts from Australia. The TWA heard that the expert from Australia would provide information to the expert from the European Community.	TWA
	An expert from the United States of America reported that, in the United States of America, distinct plants within a variety were identified according to the following definitions of "off-type" and "variant": "Variant: The term "variant" means any seed or plant which: (a) is distinct within the variety but occurs naturally in the variety; (b) is stable and predictable with a degree of reliability comparable to other varieties of the same kind, within recognized tolerances, when the variety is reproduced or reconstituted; and (c) was originally a part of the variety as released. A variant is not an off-type." "Off-type: The term "off-type" means any seed or plant not part of the variety in that it deviates in one or more characteristics from the variety as	TWC

	described and may include: a seed or plant of another variety; a seed or plant not necessarily any variety; a seed or plant resulting from cross-pollination by another kind or variety; a seed or plant resulting from uncontrolled self-pollination during production of hybrid seed; or segregates from any of the above."	

TGP/14 "Glossary of Technical, Botanical and Statistical Terms Used in UPOV Documents"

Comments on document TGP/14/1 Draft 9:

General	in the future revision (TGP/7/3), with particular regard to Section 3 "Statistical Terms", to update terms that have recently been added to TGP/14 and to delete terms that are not used in UPOV documents	TWA
	SECTION 2. BOTANICAL TERMS SUBSECTION 2. SHAPES AND STRUCTURES I. SHAPE	
1.3	to introduce the possibility to provide a different definition for the terms "base" and "apex" where that would be appropriate for the Test Guidelines concerned, in particular to avoid confusion in the use of commonly used terms by breeders. On that basis, it was agreed that the definitions of the terms should always be provided in the Test Guidelines. Furthermore, in order to ensure that applicants used the correct terms in completing the Technical Questionnaire, it was agreed that the relevant illustration of shapes in the Test Guidelines should be added to the Technical Questionnaire.	TWV
1.5	to retain the states "small" and "large" for ratio, but to add a clarification in brackets, e.g. for ratio length/width, to have "small (moderately compressed)", "large (moderately elongated)" etc.	TWV
1.5	*regarding TWV comment* the TWA agreed that it would not be appropriate to introduce the possibility to have multiple terms for the same state of expression. It recalled that Chapter 8 provided the opportunity to provide a clarification of the states of expression, whilst noting that the states should be as clear as possible for applicants in the Technical Questionnaire.	TWA
1.5 (second)	(after Chart for Other Plane Shapes) to remove reference to a decision-tree	TWV, TWA
2.10	to update cross-references	TWV, TWA
	SECTION 3 "STATISTICAL TERMS"	

<table>
<tr><td>General</td><td>To add the following introductory text:
“The definitions included in the glossary are in relation to the use of these terms in DUS examination”</td><td>TWC</td></tr>
<tr><td>Bivariate Normality</td><td>To add to following illustration:
 
</td><td>TWC</td></tr>
<tr><td>Contingency Table</td><td>to read “A contingency table is a table showing the responses of subjects to one factor as a function of another factor. For instance, the following contingency table shows a characteristic as a function of different varieties (the data are hypothetical). The entries show the number of plants for each variety with particular notes for a characteristic.
<table>
<tr><td>Characteristic
Variety</td><td>State 1</td><td>State 2</td><td>State 3</td></tr>
<tr><td>Variety A
Variety B
Variety C</td><td>18
3
6</td><td>20
10
24</td><td>2
27
10</td></tr>
</table></td><td>TWC</td></tr>
<tr><td>Random effect</td><td>To be deleted</td><td>TWC</td></tr>
<tr><td>Random Term/ Random Factor</td><td>to read “Random Term / Random Factor: A factor is random when the levels under study can be considered a random sample drawn from some large homogeneous population. A goal of the study may be to make a statement regarding the larger population. See also factor.”</td><td>TWC</td></tr>
</table>

(b) Revision of TGP Documents:

TGP/7: Development of Test Guidelines

Comments on document TGP/7/2 Draft 3

General	to replace “range of variation” with “level of variation”, or where the General Introduction is quoted, to explain that the term “level of variation” is considered to be more appropriate than the term “range of variation”, which has been used in the General Introduction (see, for example, Chapter 6.4).	TWC
	Section 1	

1.2	to explain the importance for harmonization of variety descriptions of using the Test Guidelines as individual authorities' test guidelines. In cases where that would not be possible, to encourage the inclusion of references to the characteristic number in the Test Guidelines in the individual authorities' test guidelines.	TWV, TWA
1.2	to move to the end of TGP/7	TWA
1.2.1.2	second sentence to read "Therefore, each authority may decide to request a larger quantity of plant material, for example to allow for potential losses during establishment or for a standard sample."	TWA
1.2.1.5	to clarify that the harmonization of variety descriptions could be lost if different example varieties are used in individual authorities' test guidelines	TWV
1.2.1.7	to amend to cover information provided by breeders in a breeder testing system	TWV
1.2.1.7	to explain that it may still be useful to develop a national set of example varieties in cases where example varieties are provided in the Test Guidelines or if a regional set of example varieties has been developed.	TWA
1.2.1.9	to be retained and final sentence to read "In the interim, members of the Union may indicate in DUS reports that the characteristic in the individual authorities' test guidelines has some differences to the characteristic in the Test Guidelines, pending consideration of a revision of the Test Guidelines by the Technical Committee."	TWA
1.2.1.10	to delete "including means of ensuring that applicants are aware of such changes."	TWA
1.2.1.11	to add that, according to national requirements, the authority's technical questionnaire may request additional information to that requested in the Technical Questionnaire of the UPOV Test Guidelines	TWA
2.2.4.4	to read "In advance of the TWP session, the leading expert should prepare a preliminary draft of the Test Guidelines ("Subgroup draft") for comments by the subgroup. On the basis of the comments received from the subgroup, the leading expert should establish a first draft for the TWP. This draft is sent to the Office which will produce a document for distribution to the members of the TWP(s) concerned for discussion at their session(s). Prior to the TWP session, the Office will make a preliminary check that the draft has been prepared according to document TGP/7 and, in particular, that it conforms with the TG/Template (Annex 1 *[cross ref.]*). A result of that check will be provided to the Leading Expert at least one week before the session. […]	TWV, TWA
	Annex 1: TG Template	
2.3	Netherlands to develop draft guidance on the quantity of plant material to be provided for Test Guidelines, for consideration at the forty-fourth session of the TWV with a view to its inclusion in a future revision of TGP/7 (document TGP/7/3)	TWV
4.1	to develop ASW for the assessment of distinctness of hybrids using the parental formula, on the basis of the wording in the Test Guidelines for Maize.	TWV, TWA
	Annex 2: Additional Standard Wording (ASW) for the TG Template	

ASW 8	to add ASW for assessment of uniformity of ear-row / panicle row plots as follows: "For the assessment of uniformity of [plants, parts of plants] / [ear-rows] / [panicle-rows], a population standard of { x }% and an acceptance probability of at least { y } % should be applied. In the case of a sample size of { a } [plants, parts of plants] / [ear-rows] / [panicle rows], [{ b } off-types [plants, parts of plants] / [ear-rows] / [panicle-rows] are] / [1 off-type [ear-row] / [panicle-row] is] allowed." "[An ear-row] / [A panicle-row] is considered to be an off-type [ear-row] / [panicle-row] if there is more than one off-type plant within that [ear-row] / [panicle-row]"	TWA
ASW 8	to develop ASW for specific characteristics that might be observed on different sample sizes	TWA
ASW 8	to introduce following ASW for hybrid varieties where parental formula used: "Where the assessment of distinctness of hybrids involves a pre-screening system on the basis of the parental lines and formula, the uniformity of a hybrid variety should, in addition to an examination of the hybrid variety itself, also be assessed by examination of the uniformity of its parent lines."	TWA
ASW 13	to include an indication that the parental formula would be used	TWV, TWA
ASW 15	to delete version (b) and move ASW 15(a) to the TG Template	TWA
	Annex 3: Guidance Notes (GN) for the TG Template	
GN 9	to add the ISBN number for the "Growth stages of mono-and dicotyledonous plants - BBCH Monograph"	TWA
GN 19 (3)	to add an example to clarify the meaning	TWA
GN 20 (3.1)	to complete all states, i.e. including the even states, in the "length of stem" example	TWA
GN 20 (3.7)	to delete Example 1	TWA
GN 26	to explain that it is more appropriate to use the chronological order if groups of characteristics are to be observed at the same time	TWA

GN 28	the TWV noted that it would not be able to review any proposed amendments to GN 28 before the Technical Committee considered the approval of document TGP/7/2 in 2010. The TWV noted the importance of example varieties in Test Guidelines for vegetable crops and generally supported the text in GN 28. Therefore, to avoid a delay in the adoption of document TGP/7/2, it proposed that document TGP/7/2 should be adopted in 2010 without amendments to GN 28 and that any proposed amendments should be considered in a future revision of document TGP/7, if appropriate. TWA: agreed and also agreed to add an agenda item to discuss example varieties at its thirty ninth session	TWV, TWA
GN 31	to add the possibility to indicate that the variety is a parent line, with a reference to document TGP/5 "Experience and Cooperation in DUS Testing", Section 11/1 "Examples of Policies and Contracts for Material Submitted by the Breeder", which explains in paragraph 1.1 that "[…] in the particular case of parent lines submitted as a part of the examination of a candidate hybrid variety, living plant material should only be made available to other variety collectors in such a way that the legitimate interests of the breeder would be safeguarded."	TWV
GN 31	*regarding TWV comment:* the TWA did not agree that it was necessary to make such an indication in the Technical Questionnaire for a parent line submitted as a part of an application for a hybrid variety because the information concerning such parent lines would be included in a single application for the hybrid variety.	TWA
GN 32	Three-Way Hybrid: to add a line to enter the name of the female hybrid parent	TWV, TWA
Annex 4: Collection of approved characteristics		
General	the TWA noted that document TWA/38/3, paragraph 31, explained that the "TC noted that the Office of the Union planned to develop an improved TG Template and to integrate the Collection of Approved Characteristics into that template in a user-friendly package for drafters of Test Guidelines.". It heard that the experience of the Office of the Union was that the collection of approved characteristics was not, in general, used by Leading Experts in the drafting of Test Guidelines and agreed that it would not be a good use of resources to invest a substantial effort in its development for the time-being.	TWA

[Appendix to Annex II follows]

APPENDIX to ANNEX II

Proposed new text for TGP/8/1 Draft 13, Part II, Section 5

5. MATCH METHOD[1]

5.1 Requirements for application of method

5.1.1 The match method is appropriate for assessing distinctness of varieties where:

- data from more than one year are analyzed,
- observations made on a plant (or plot) in the second year are compared to observations made by the breeder in the first year,
- there are claimed differences between plants (or plots) of a variety based on information from the first year trial,
- the requirements of the method depend on the particular statistical test that is used (e.g. LSD, Multiple Range Tests (MRT), Chi-square or Fisher's Exact).

5.2 Match Method

5.2.1 The Match method to assess distinctness was developed for use where the trials are conducted by the breeder in the first year and examined by the testing authority in the second year (see document TGP/6 section 2/1). Whether differences are sufficiently consistent is assessed using a statistical test (e.g. LSD, MRT, Chi-Square or Fisher's Exact) to gauge whether the differences in the second year are significant and agree with the "direction of the differences" declared by the breeders in the first year. The choice of statistical test depends on the type of expression of the characteristic concerned. For two varieties to be distinct using the Match method, the varieties need to be significantly different in the same direction claimed by the breeder in the first year.

5.2.2 The requirements of the method depend on the particular statistical test that is used (e.g. LSD, MRT, Chi-Square or Fisher's Exact). For quantitative characteristics the statistical test may for example be based on a one-tailed LSD, if there is one candidate, or on a one-tailed MRT, if there is more than one candidate included in the growing trial. A Chi-square test or Fisher's exact test may be used for pseudo-qualitative or qualitative characteristics where the requirements for these tests are met. Although these tests are most useful in trials of cross-pollinated varieties, they can be similarly applied to trials of self pollinated and vegetatively propagated varieties provided the relevant requirements are met

[1] Comment by expert from Australia: Discussion of ordinal and nominal scales have been moved out of TGP/8 for further development (see TWC27/11). It is agreed that paragraphs 5.1 and 5.2 do not sit well here as they are applicable to nominal data generally. If they are to be removed they should be placed elsewhere in TGP/8 as they place use of this type of data in context.

5.2.3 The Match method typically involves relatively small scale trials where the number of candidate and reference varieties in the trials is limited to the most similar varieties of common knowledge.

5.3 Pearson's chi-square test applied to contingency tables

5.3.1 A contingency table is a table showing the responses of subjects to one factor as a function of another factor. In DUS testing it is generally used for categorical data where individuals of a variety can be allocated to discrete states of expression for a characteristic. Various statistical tests can be used to analyze the data in contingency tables depending on the particular circumstances. For example, Pearson's Chi-square test, as applied to contingency tables, is useful where:

- observations on a characteristic are allocated to two or more categories (classes) and are recorded in a contingency table
- there are some differences between plants (or plots) of a variety;
- the only source of variation should be caused by random sampling, e.g. there should be no variation due to soil conditions, etc[2]
- the minimum expected value in each category should be five

5.3.2 In some cases, distinctness may be established by classifying individual varieties into broad groups and demonstrating statistically different grouping patterns for different varieties. Examples include counts based on broad flower color groups - such as dark blue violet versus not dark blue violet and the disease/pest/nematode infection classes. Data based on counts of individuals in a sample/population belonging to each of several classes require statistical analysis capable of dealing with categorical data.

5.3.3 To use the Chi-square analysis for plant breeder rights' (PBR) purposes, we should consider how we are going to arrive at certain conclusions about distinctness by formulating certain hypotheses using the classification data.

The standard formula for the chi-square statistic used in such analysis is:

$$\chi^2 = \Sigma \frac{(\text{Observed value of a class} - \text{Expected value of a class})^2}{\text{Expected value}}$$

5.3.4 Hence, the Chi-square distribution is a continuous distribution based upon an underlying normal distribution.

5.3.5 The following precautions are to be considered before using the chi-square test.

(1) Selection of the hypothesis to be tested should be based on previously known facts or principles

[2] Comment by expert from Australia: – is this covered in Part 1? Sources of variation are discussed in section 1.5.3.3. Particularly 1.5.3.3.4. Also in 1.5.3.3.7.3.1. TGP/8 does not specifically refer to random sampling although the term is defined in TGP/14. Propose that this element is removed as 'limiting other sources of variation' should be covered generally elsewhere as the requirement applies equally to the other methods.

(2) Given the hypothesis, you should be able to assign expected values for each class correctly. Avoid using the chi-square test if the smallest expected class is less than five. By increasing the sample size the size of the smallest expected class can be made larger. Alternatively, if some classes have a size less than five, either pool those adjacent classes to bring the size of the pooled class to five or more than five, or use an exact test.

(3) Degrees of freedom is defined as the number of classes that are independent to be assigned an arbitrary value. For example, if we have two classes the degrees of freedom is 2-1 = 1. Hence, in using this method to test a hypothesis, the degrees of freedom for the chi-square test is one less than the number of classes.

(4) Avoid using two class situations which follow more like the binomial distribution, with np or nq less than 5. If you encounter such situations, calculate expected values using formulae based on the binomial distribution. In a two class situation, np is the size of one of the classes determined by the number of events (n) times the probability of falling into that class (p). Similarly the size of the other class (nq) is determined by n times the probability (q) of falling into that class. So in a situation where the probability of falling into either class is equal (p=q=0.5) and the sample size is 10 (n) the number expected in each class is 5. Always use Yates Correction for determining the chi-square test with only one degree of freedom.

5.3.6 Let us examine the following data on the disease scoring of a Lucerne candidate variety and its four reference varieties. The disease scored was *Colletotrichum trifolii* (Characteristic 19, TG/6/5, Lucerne). The scoring was on a 5 class scale, with class 1 (note 9) being resistant and class 5 (note 1) being susceptible.

Contingency table of number of plants counted in different classes in each variety after 7-10 days of inoculation

Note(Class)	Candidate		Reference 1	Reference 2	Reference 3	Reference 4
9(1)	34		12	6	1	7
7(2)	4		7	6	5	10
5(3)	1		9	5	5	5
3(4)	1		7	9	8	7
1(5)	6		9	19	9	15
Total	46		44	45	28	44

5.3.7 It can be seen from the table that the candidate variety has more plants in the resistant category than the reference varieties. However, to statistically test the significance of the difference, we need to formulate a hypothesis:

(1) Whether the reference varieties differ significantly or not from the candidate in the distribution of scores i.e. by testing the null hypothesis. The null hypothesis in this case is all the varieties show similar reaction to the Colletotrichum crown rot. This can be done by testing the "distinctness X^2".

5.3.8 Pooling of classes to form a new intermediary pooled class is necessary to meet the minimum expected value requirement for the use of the chi square test.

Now the observed data is reduced to:

Class/Score	Candidate		Reference variety 1	Reference variety 2	Reference variety 3	Reference variety 4
1	*34*		12	6	1	7
2	*6*		23	20	18	22
3	*6*		9	19	9	15
Total	*46*		44	45	28	44

5.3.9 For each comparison of the candidate with each reference variety, a two-way table of observed values is formed. The expected values are calculated as the product of the row and column totals divided by the grand total, and the chi square statistic is calculated. The distributions of expected values for different varieties are as follows:

Observed for Reference Variety 1

Class/Score	**Candidate**	**Reference variety 1**	**Total**
1	34	12	46
2	6	23	29
3	6	9	15
Total	**46**	**44**	**90**

Expected for Reference Variety 1

Class/Score	**Candidate**	**Reference variety 1**	**Total**
1	23.5=46x46/90	22.5=46x44/90	46
2	14.8=29x46/90	14.2=29x44/90	29
3	7.7=15x46/90	7.3=15x44/90	15
Total	46	44	90

Similarly, using the table of observed data in 5.3.8, the expected values for reference varieties 2,3 and 4 are;

Class/Score	**Candidate**	**Reference variety 2**	**Total**
1	20.2	19.8	40
2	13.1	12.9	26
3	12.6	12.4	25
Total	46	45	91

Class/Score	Candidate	Reference variety3	Total
1	21.8	13.2	35
2	14.9	9.1	24
3	9.3	5.7	15
Total	46	28	74

Class/Score	Candidate	Reference variety4	Total
1	21.0	20.0	41
2	14.3	13.7	28
3	10.7	10.3	21
Total	46	44	90

~~5.3.10~~[3]

5.3.11 For calculating the "distinctness X^2" for Reference variety 1

X^2 = (34-23.5)2/23.5 + (12-22.5)2/22.5 + (6-14.8)2/14.8 + (23-14.2)2/14.2 + (6-7.7)2/7.7 + (9-7.3)2/7.3
= 21.1
on (No rows – 1)(No cols – 1) = 2 df

5.3.12 The number of degrees of freedom for looking up the χ^2 table is one less than the number of rows multiplied by one less than the number of columns i.e., 3 – 1 x 2-1 =2.

5.3.13 At P = 0.01, for 2 df, the tabular value is 9.21. The calculated distinctness X^2 is more than the tabulated χ^2 value. Therefore, we reject the null hypothesis that Reference variety 1 has a similar reaction to the disease as the candidate variety.

5.3.14 Similarly the calculated "distinctness X^2" for Reference variety-2, Reference variety-3 and Reference variety-4 are 33.9, 35.4 and 30.8, respectively, which are all greater than the tabulated $\chi 2$ value of 9.21 at 2 df.

5.3.15 Hence, all the Reference varieties are significantly different from the candidate variety in reaction to Colletotrichum crown rot.

[End of Annex II and of document]

[3] Comment by expert from Australia: Deletion of 5.3.10 is a necessary consequence of the deletion in 5.3.9 as it relates directly to the information deleted. Also 5.3.12 should now be retained as it is the only place that shows how to look up the value from the Chi table.