



TC/51/16
 ORIGINAL : anglais
 DATE : 11 février 2015

UNION INTERNATIONALE POUR LA PROTECTION DES OBTENTIONS VÉGÉTALES

Genève

COMITE TECHNIQUE

Cinquante et unième session Genève, 23–25 mars 2015

RÉVISION DU DOCUMENT TGP/8 :
 PREMIERE PARTIE : PROTOCOLE D'ESSAI DHS ET ANALYSE DES DONNEES,
 NOUVELLE SECTION : RÉDUCTION OPTIMALE DE LA VARIATION DUE
 À DIFFÉRENTS OBSERVATEURS

Document établi par le Bureau de l'Union

Avertissement : le présent document ne représente pas les principes ou les orientations de l'UPOV

1. L'objet du présent document est de présenter le projet d'une nouvelle section pour le document TGP/8 : première partie : Protocole d'essai DHS et analyse des données, qui serait intitulée : "Réduction optimale de la variation due à différents observateurs".

2. Les abréviations ci-après sont utilisées dans le présent document :

CAJ :	Comité administratif et juridique
TC :	Comité technique
TC-EDC :	Comité de rédaction élargi
TWA :	Groupe de travail technique sur les plantes agricoles
TWV :	Groupe de travail technique sur les plantes potagères
TWC :	Groupe de travail technique sur les systèmes d'automatisation et les programmes d'ordinateur
TWF :	Groupe de travail technique sur les plantes fruitières
TWO :	Groupe de travail technique sur les plantes ornementales et les arbres forestiers
TWP :	Groupes de travail techniques

3. Le présent document est structuré comme suit :

INFORMATIONS GENERALES	2
FAITS NOUVEAUX EN 2014	2
COMITE TECHNIQUE	2
GROUPES DE TRAVAIL TECHNIQUES	3

ANNEXE : PROJET DE CONSEILS À INCORPORER DANS UNE FUTURE VERSION RÉVISÉE DU DOCUMENT TGP/8 CONCERNANT LA RÉDUCTION OPTIMALE DE LA VARIATION DUE À DIFFÉRENTS OBSERVATEURS

INFORMATIONS GENERALES

4. Il est indiqué au paragraphe 2.9.1 “Contrôle de la variation due à différents observateurs” du document TGP/8/1, projet 7, première partie, examiné par les groupes de travail techniques à leurs sessions de 2007 :

“[Si cette section est requise, les TWP sont invités à formuler des conseils sur le contrôle de la variation due à différents observateurs lorsque l’analyse statistique n’est pas utilisée pour déterminer la distinction, et à l’examiner en liaison avec le paragraphe 2.7.2.9.]”

5. Les faits nouveaux survenus avant 2014 sont présentés dans le document TC/50/21 “Révision du document TGP/8 : première partie : Protocole d’essai DHS et analyse des données, nouvelle section : Réduction optimale de la variation due à différents observateurs”.

FAITS NOUVEAUX EN 2014

Comité technique

6. À sa cinquantième session tenue à Genève du 7 au 9 avril 2014, le TC a examiné le document TC/50/21 “Révision du document TGP/8 : première partie : Protocole d’essai DHS et analyse des données, nouvelle section: Réduction optimale de la variation due à différents observateurs”.

7. Le TC a pris note que le TWF avait demandé à un expert de la Nouvelle-Zélande de rendre compte, à sa session de 2014, des travaux réalisés précédemment sur la description variétale harmonisée de la pomme pour un ensemble convenu de variétés (voir le paragraphe 46 du document TC/50/36 “Compte rendu des conclusions”).

8. Le TC a invité l’expert de l’Australie à rédiger, avec l’aide d’experts de l’Union européenne, de la France, de l’Allemagne, des Pays-Bas et du Royaume-Uni, des conseils supplémentaires en vue de leur incorporation dans une future version révisée du document TGP/8, concernant l’optimisation de la variation due à différents observateurs, y compris des conseils sur les caractères PQ et QN/MG, afin que ces conseils soient examinés par les TWP lors de leurs sessions de 2014 (voir le paragraphe 47 du document TC/50/36 “Compte rendu des conclusions”).

9. En réponse à la demande du TC, le rédacteur de l’Australie (M. Nik Hulse) a rédigé, avec l’aide d’experts de l’Union européenne, de la France, de l’Allemagne, des Pays-Bas et du Royaume-Uni, des conseils supplémentaires en vue de leur incorporation dans une future version révisée du document TGP/8, concernant l’optimisation de la variation due à différents observateurs, y compris des conseils sur les caractères PQ et QN/MG, afin que ces conseils soient examinés par les TWP lors de leurs sessions de 2014.

10. L’expert de l’Australie a proposé que les TWP examinent les points ci-après en vue de les inclure, ou d’inclure des précisions, dans le prochain projet de conseils à incorporer dans une future version révisée du document TGP/8 concernant l’optimisation de la variation due à différents observateurs :

- Comme déjà indiqué dans le projet, les caractères QN/MG peuvent être traités de la même façon que les caractères QN/MS. Toutefois, il convient de noter que dans le cas des caractères QN/MG, il faut également tenir compte de l’effet que peut avoir une variation aléatoire à l’intérieur d’une parcelle.
- Il convient de déterminer plus précisément comment inclure des conseils sur les caractères PQ. Il faudrait indiquer que le document TGP/14 apporte également des précisions utiles au sujet de nombreux caractères PQ (p. ex. en ce qui concerne la forme). Les différences entre observateurs pourraient éventuellement être mesurées sur la base de la fréquence des écarts.
- Actuellement, le document met l’accent sur la variation entre observateurs au niveau du service. Il serait utile de déterminer s’il conviendrait de mentionner la question de l’optimisation de la variation due à différents observateurs dans ce document ou éventuellement dans un nouveau document. À noter que cela impliquerait un plus grand nombre de facteurs tels que les interactions génotype-environnement, les méthodes d’échantillonnage et les essais d’étalonnage.

Groupes de travail techniques

11. Le TWO, le TWF, le TWC, le TWV et le TWA ont examiné les documents TWO/47/15, TWF/45/15, TWC/32/15, TWV/48/15 et TWA/43/15 "Revision of Document TGP/8 : Part I : DUS Trial Design and Data Analysis, New Section : Minimizing the Variation due to Different Observers", respectivement (voir les paragraphes 39 à 41 du document TWO/47/28 "Report", les paragraphes 26 à 32 du document TWF/45/32 "Report", les paragraphes 50 à 53 du document TWC/32/28 "Report", les paragraphes 34 à 37 du document TWV/48/43 "Report" et les paragraphes 28 à 33 du document TWA/43/27 "Report").

12. Le TWO a pris note que le TWF avait demandé à un expert de la Nouvelle-Zélande de rendre compte, à sa session de 2014, des travaux réalisés précédemment sur la description variétale harmonisée de la pomme pour un ensemble convenu de variétés, ainsi qu'indiqué au paragraphe 18 du document TWO/47/15 (voir le paragraphe 40 du document TWO/47/28 "Report").

13. Le TWO, le TWF, le TWC, le TWV et le TWA sont convenus que le projet de conseils présenté à leurs sessions à l'annexe des documents TWO/47/15, TWF/45/15, TWC/32/15, TWV/48/15 et TWA/43/15, respectivement, devrait être élaboré plus avant en vue de son incorporation dans une future version révisée du document TGP/8, concernant l'optimisation de la variation due à différents observateurs, y compris des conseils sur les caractères PQ et QN/MG, compte tenu des points soulevés par l'expert de l'Australie au paragraphe 21 des documents TWO/47/15, TWF/45/15, TWC/32/15, TWV/48/15 et TWA/43/15 (voir le paragraphe 41 du document TWO/47/28 "Report", le paragraphe 27 du document TWF/45/32 "Report", le paragraphe 52 du document TWC/32/28 "Report", le paragraphe 34 du document TWV/48/43 "Report" et le paragraphe 30 du document TWA/43/27 "Report").

14. Le TWO et le TWA sont convenus que le document devrait mettre l'accent sur la variation entre observateurs au niveau du service et non pas sur l'optimisation de la variation entre services (voir le paragraphe 41 du document TWO/47/28 "Report" et le paragraphe 31 du document TWA/43/27 "Report").

15. Le TWF a suivi un exposé présenté par les experts de l'Allemagne et de la Nouvelle-Zélande sur les travaux réalisés précédemment sur la description variétale harmonisée de la pomme pour un ensemble convenu de variétés, qui figure dans le document TWF/45/28 (voir le paragraphe 28 du document TWF/45/32 "Report").

16. Le TWF est convenu de l'importance de l'optimisation de la variation entre observateurs au niveau du service et également entre service et a donc suggéré d'examiner la possibilité de mener une étude sur un nouveau projet sur la description variétale harmonisée pour un ensemble convenu de variétés. L'expert de l'Allemagne a proposé de présenter au TWF, à sa quarante-sixième session, un protocole de projet ainsi qu'une liste convenue de variétés à examiner, pour déterminer s'il serait judicieux d'approfondir l'étude (voir le paragraphe 30 du document TWF/45/32 "Report").

17. Le TWF et le TWV ont souligné l'importance de la qualité des principes directeurs d'examen dans la fourniture d'orientations claires aux examinateurs DHS et la garantie de la cohérence des observations entre les observateurs de chaque service, ainsi que l'importance de la formation continue pour les examinateurs (voir le paragraphe 31 du document TWF/45/32 "Report" et le paragraphe 35 du document TWV/48/43 "Report").

18. Le TWF a invité l'expert de l'Australie à rendre compte, à sa quarante-sixième session, de l'incidence que pouvaient avoir l'endroit, l'observateur et l'année sur la conformité d'un caractère pour une plante spécifique (voir le paragraphe 32 du document TWF/45/32 "Report").

19. Le TWC a pris note que l'expert de la Nouvelle-Zélande avait rendu compte à la cinquante-cinquième session du TWF des travaux réalisés précédemment sur la description variétale harmonisée de la pomme pour un ensemble convenu de variétés (voir le paragraphe 51 du document TWC/32/28 "Report").

20. Le TWC est convenu que le projet de conseils devrait commencer par la variation entre observateurs au niveau du service, puis entre différents services (voir le paragraphe 53 du document TWC/32/28 "Report").

21. Le TWV a suggéré d'inclure un exercice de formation dans un essai DHS comme base pour le partage de données d'expérience dans ce domaine et d'exploiter davantage le TWV à des fins de formation (voir le paragraphe 36 du document TWV/48/43 "Report").

22. Le TWA a pris note qu'un expert de la Nouvelle-Zélande avait présenté un exposé à la quarante-cinquième session du TWF sur les travaux réalisés précédemment sur la description variétale harmonisée de la pomme pour un ensemble convenu de variétés, qui figure dans le document TWF/45/28 "Harmonized example varieties for Apple: historical data and possible new developments" (voir le paragraphe 29 du document TWA/43/27 "Report").

23. Le TWA a souligné l'importance de la qualité des principes directeurs d'examen dans la fourniture d'orientations claires aux examinateurs DHS et la garantie de la cohérence des observations, ainsi que l'importance de la formation continue pour les examinateurs. Le TWA est convenu de proposer une recommandation générale selon laquelle, dans la mesure du possible, un essai devrait être confié à un même observateur pour éviter toute variation dans les observations (voir le paragraphe 32 du document TWA/43/27 "Report").

24. Le TWA est convenu que les caractères QN/MG pouvaient être traités de la même façon que les caractères QN/MS et a noté qu'il fallait également tenir compte de l'effet que pouvait avoir une variation aléatoire à l'intérieur d'une parcelle. Le TWA est convenu que les différences entre observateurs concernant les caractères PQ pourraient être examinées au moyen de méthodes non paramétriques, par exemple la fréquence des écarts (voir le paragraphe 33 du document TWA/43/27 "Report").

25. En réponse aux observations formulées par les TWP à leurs sessions en 2015, M. Nik Hulse (Australie) a élaboré un nouveau projet de conseils destiné à être intégré dans le document TGP/8 : première partie : Protocole d'essai DHS et analyse des données, nouvelle section : Réduction optimale de la variation due à différents observateurs, qui figure à l'annexe du présent document, afin que les TWP examinent ces conseils lors de leurs sessions de 2015.

26. *Le TC est invité*

a) à examiner le projet de conseils destiné à être intégré dans le document TGP/8, concernant l'optimisation de la variation due à différents observateurs, qui figure à l'annexe du présent document, compte tenu des observations formulées par les TWP à leurs sessions de 2014; et

b) à demander aux TWP d'examiner à leurs sessions de 2015 un nouveau projet contenant les modifications éventuelles convenues par le TC à sa cinquante et unième session.

[L'annexe suit]

DOCUMENT TGP/8/1 : PREMIÈRE PARTIE : NOUVELLE SECTION :
OPTIMISATION DE LA VARIATION DUE À DIFFÉRENTS OBSERVATEURS DANS UN MÊME ESSAI

1. Introduction

Le présent document concerne les caractères QN/MG, QN/MS, QN/VG et QN/VS. Il ne traite pas expressément des caractères PQ tels que la couleur ou la forme. La méthode Kappa décrite est dans une large mesure applicable à ces caractères, le caractère Kappa standard étant prévu par exemple pour les données nominales. Toutefois, à notre connaissance, cette méthode n'a pas été appliquée à des caractères PQ, ces dernières pouvant également nécessiter des informations supplémentaires en matière d'étalonnage. À titre d'exemple, pour l'étalonnage des couleurs, il faut également tenir compte du code de couleur RHS, des conditions d'éclairage, etc. Ces éléments ne sont pas traités dans le présent document.

1.1 La variation des mesures ou des observations peut s'expliquer par de nombreux facteurs différents tels que le type de culture, le type de caractère, l'année, le site, le protocole et la gestion des essais, la méthode et l'observateur. Notamment en ce qui concerne les caractères observés visuellement (QN/VG ou QN/VS), les différences d'un observateur à l'autre peuvent expliquer de fortes variations et un biais éventuel entre observations. Il se peut qu'un observateur soit moins bien formé ou interprète d'une manière différente le caractère. Par conséquent, lorsque l'observateur A mesure la variété 1 et l'observateur B la variété 2, la différence mesurée peut s'expliquer par des différences entre observateurs A et B au lieu de différences entre variétés 1 et 2. Il ne fait aucun doute que nous nous intéressons principalement aux différences entre variétés et non aux différences entre observateurs. Il est important de comprendre que la variation due aux différents observateurs ne peut pas être supprimée mais qu'il existe des moyens de la contrôler.

1.2 Il est recommandé, dans la mesure du possible, de confier l'essai à un même observateur pour réduire de façon optimale la variation due à différents observateurs.

2. Formation et importance de fournir des indications claires sur les caractères et la méthode d'observation

2.1 Il est essentiel de former les nouveaux observateurs pour des raisons d'harmonisation et de continuité des observations. Les manuels d'étalonnage, la supervision et les conseils des observateurs expérimentés, de même que l'utilisation de variétés indiquées à titre d'exemple pour illustrer la gamme d'expression, sont autant de moyens à exploiter pour y parvenir.

2.2 Les principes directeurs d'examen de l'UPOV visent à harmoniser le processus de description variétale et à décrire aussi clairement que possible les caractères d'une plante et leurs niveaux d'expression. Il s'agit de la première étape de contrôle de la variation et du biais. Toutefois, la façon dont un caractère est observé ou mesuré peut varier en fonction du lieu ou du service chargé de l'examen. Les manuels d'étalonnage mis au point par le service local chargé de l'examen sont très utiles pour l'adaptation locale des principes directeurs d'examen de l'UPOV. Le cas échéant, ces manuels propres à une plante donnée expliquent plus en détail les caractères à observer, et précisent le lieu et les modalités d'observation. En outre, ils peuvent contenir des photos et des dessins pour chaque caractère, souvent pour chaque niveau d'expression d'un caractère.

2.4 Le Glossaire des termes utilisés dans les documents de l'UPOV (document TGP/14/2) donne des précisions utiles sur de nombreux caractères, notamment les caractères PQ.

2.3 Après que l'observateur a été formé, il est important de veiller à ce qu'il bénéficie régulièrement de cours de perfectionnement et passe un test d'étalonnage.

3. Essai d'étalonnage

3.1 Après la formation d'un observateur, l'étape suivante consiste à vérifier les acquisitions des observateurs dans le cadre d'un essai d'étalonnage. Cela est particulièrement utile pour les observateurs inexpérimentés qui doivent procéder à des observations visuelles (caractères QN/VG et QN/VS). S'ils se livrent à des observations visuelles, ils devraient de préférence faire un essai d'étalonnage avant de formuler des observations durant l'essai. Mais il est aussi important que les observateurs expérimentés vérifient eux-mêmes leurs connaissances régulièrement pour s'assurer qu'ils satisfont toujours aux critères d'étalonnage.

3.2 Un essai d'étalonnage peut être organisé et analysé de différentes manières. En général, il consiste à faire mesurer les mêmes éléments par plusieurs observateurs avant d'évaluer les différences entre eux.

4. Essai d'étalonnage pour les caractères QN/MG ou QN/MS

4.1 Pour les observations effectuées à l'aide d'instruments de mesure, tels que des règles (généralement pour les caractères QN/MS), la mesure est souvent effectuée à l'aide d'une échelle d'intervalle ou de rapport. Dans ce cas, la méthode de Bland et Altman (1986) peut être utilisée. Cette méthode consiste à faire figurer les résultats obtenus par paire d'observateurs dans un diagramme de dispersion et à les comparer avec la ligne $y=x$. Cela permet d'évaluer à l'œil le degré de corrélation des mesures. L'étape suivante consiste à prendre la différence par objet et à construire un graphique avec, sur l'axe y, la différence entre les observateurs et, sur l'axe x, l'indice ou la valeur moyenne de l'objet. En traçant les lignes horizontales $y=0$, $y=\text{moyenne}(\text{dif})$ et les deux lignes $y = \text{moyenne}(\text{dif}) \pm 2 \times \text{l'écart type}$, le biais entre observateurs et toute autre valeur atypique peuvent être facilement détectés. De même, on peut également étudier la différence entre la mesure de chaque observateur et la moyenne de toutes les mesures des observateurs. Des méthodes d'essai telles que le test t par paire peuvent être utilisées en cas de différence importante d'un observateur à l'autre ou par rapport à la moyenne des autres observateurs.

4.2 En prenant deux mesures chez chaque observateur pour chaque objet, on peut observer les différences entre ces deux mesures. Si ces différences sont importantes par rapport à celles d'autres observateurs, cet observateur pourrait avoir une faible reproductibilité. Si l'on compte pour chaque observateur le nombre de valeurs aberrantes modérées et élevées (p. ex. supérieures à 2 et 3 fois l'écart type respectivement), on obtient un tableau indiquant le nombre de valeurs aberrantes par observateur qui peut être utilisé pour déterminer si l'observateur satisfait aux exigences d'assurance-qualité.

4.3 Les contrôles de la qualité peuvent également être fondés sur des essais de reproductibilité ou de répétabilité pour laboratoires tels que ceux décrits dans la norme ISO 5725-2. Un logiciel gratuit disponible sur le site Web de l'ISTA permet d'obtenir des valeurs et des graphiques en accord avec cette norme ISO.

4.4 Dans de nombreux cas de QN/MG ou QN/MS, des instructions claires sont en général suffisantes et la variation ou le biais dans les mesures entre observateurs est souvent négligeable. En cas de doute, des essais d'étalonnage tels que décrits ci-dessus peuvent contribuer à préciser les choses.

4.5 Dans le cas de caractères QN/MG, il faut tenir compte de l'effet que peut avoir une variation à l'intérieur d'une parcelle.

5. Essai d'étalonnage pour les caractères QN/VS ou QN/VG

5.1 Aux fins de l'analyse des données ordinales (caractères QN/VS ou QN/VG), il est très instructif d'établir des tableaux de contingence par paire d'observateurs pour les résultats différents. Le test de comparaison des données de Wilcoxon permet de mettre en évidence une différence structurelle (biais) entre deux observateurs.

5.2 Pour mesurer le degré de concordance, on recourt souvent au test du Kappa de Cohen (Cohen, 1960). Ce test statistique s'efforce de rendre compte des concordances aléatoires : $\kappa = P(\text{concordance}) - P(e) / (1 - P(e))$, où $P(\text{concordance})$ est la fraction d'objets se situant dans la même classe pour les deux observateurs (la diagonale principale du tableau de contingence) et $P(e)$ la probabilité de concordance aléatoire, compte tenu des valeurs marginales (comme dans le test du Khi carré). Lorsque les observateurs sont en parfaite concordance, valeur Kappa $\kappa = 1$. S'il n'y a pas de concordance entre les observateurs au-delà des résultats pouvant être le fruit du hasard ($P(e)$), $\kappa = 0$.

5.3 Le test statistique Kappa de Cohen standard envisage uniquement la possibilité d'une concordance parfaite ou d'une discordance. Si l'on met en évidence le degré de discordance (par exemple, avec les caractères ordinaux), on peut appliquer un test du Kappa linéaire ou carré (Cohen, 1968). Si l'on souhaite avoir une statistique unique pour tous les observateurs simultanément, on peut calculer un coefficient Kappa général. La plupart des programmes statistiques, dont SPSS, Genstat et R (programme Concord), fournissent des instruments de calcul du coefficient Kappa.

5.4 Ainsi qu'il a été noté, une valeur κ faible indique une faible concordance alors que les valeurs avoisinant 1 indiquent une excellente concordance. Souvent, les chiffres entre 0,6 et 0,8 sont réputés dénoter une concordance importante et les chiffres au-dessus de 0,8 une concordance presque parfaite. Si nécessaire, il existe des écarts réduits pour Kappa (à supposer que la distribution est à peu près normale). Les critères appliqués aux experts DHS chevronnés pourraient être plus stricts que ceux appliqués aux personnes inexpérimentées.

6 Protocole d'essai

6.1 En cas de pluralité d'observateurs dans un essai, la meilleure méthode consiste à confier à une même personne l'observation d'une ou plusieurs répétitions complètes. Dans ce cas, la correction des effets de blocs tient aussi compte du biais entre observateurs. Lorsqu'il est nécessaire de recourir à plus d'un observateur par répétition, il convient d'accorder davantage d'attention à l'étalonnage et à la concordance. Dans certains cas, des plans en blocs incomplets (plans alpha) peuvent être utiles, et un observateur peut être affecté aux sous-blocs. On peut ainsi corriger les différences systématiques entre observateurs.

7. Kappa de Cohen : exemple

7.1 Dans cet exemple, il y a trois observateurs et 30 objets (parcelles ou variétés). Le caractère est observé sur une échelle de 1 à 6. Les données brutes et leurs notes figurent dans les tableaux qui suivent.

Variété	Observateur		
	1	2	3
V1	1	1	1
V2	2	1	2
V3	2	2	2
V4	2	1	2
V5	2	1	2
V6	2	1	2
V7	2	2	2
V8	2	1	2
V9	2	1	2
V10	3	1	3
V11	3	1	3
V12	3	2	2
V13	4	5	4
V14	2	1	1
V15	2	1	2
V16	2	2	3
V17	5	4	5
V18	2	2	3
V19	1	1	1
V20	2	2	2
V21	2	1	2
V22	1	1	1
V23	6	3	6
V24	5	6	6
V25	2	1	2
V26	6	6	6
V27	2	6	2
V28	5	6	5
V29	6	6	5
V30	4	4	4

Notes de la variété pour	Observateur					
	1	2	3	4	5	6
V1	3	0	0	0	0	0
V2	1	2	0	0	0	0
V3	0	3	0	0	0	0
V4	1	2	0	0	0	0
V5	1	2	0	0	0	0
V6	1	2	0	0	0	0
V7	0	3	0	0	0	0
V8	1	2	0	0	0	0
V9	1	2	0	0	0	0
V10	1	0	2	0	0	0
V11	1	0	2	0	0	0
V12	0	2	1	0	0	0
V13	0	0	0	2	1	0
V14	2	1	0	0	0	0
V15	1	2	0	0	0	0
V16	0	2	1	0	0	0
V17	0	0	0	1	2	0
V18	0	2	1	0	0	0
V19	3	0	0	0	0	0
V20	0	3	0	0	0	0
V21	1	2	0	0	0	0
V22	3	0	0	0	0	0
V23	0	0	1	0	0	2
V24	0	0	0	0	1	2
V25	1	2	0	0	0	0
V26	0	0	0	0	0	3
V27	0	2	0	0	0	1
V28	0	0	0	0	2	1
V29	0	0	0	0	1	2
V30	0	0	0	3	0	0

Le tableau de contingence pour les observateurs 1 et 2 est le suivant :

O1\O2	1	2	3	4	5	6	Total
1	3	0	0	0	0	0	3
2	10	5	0	1	0	1	17
3	2	1	0	0	0	0	3
4	0	0	0	1	0	0	1
5	0	0	0	1	0	2	3
6	0	0	1	0	0	2	3
Total	15	6	1	3	0	5	30

Le coefficient Kappa entre l'observateur 1 et l'observateur 2, $\kappa(O1,O2)$, est calculé comme suit :

- $\kappa(O1,O2) = P(\text{concordance entre } O1 \text{ et } O2) - P(e) / (1 - P(e))$, où :
- $P(\text{concordance}) = (3+5+0+1+0+2)/30 = 11/30 \approx 0.3667$ (éléments diagonaux)
- $P(e) = (3/30).(15/30) + (17/30).(6/30) + (3/30).(1/30) + (1/30).(3/30) + (3/30).(0/30) + (3/30).(5/30) \approx 0.1867$. (marges par paire)
- Donc $\kappa(O1,O2) \approx (0.3667-0.1867) / (1-0.1867) \approx 0.22$

Il s'agit d'une valeur faible, indiquant une concordance très faible entre ces deux observateurs. Cela constitue un sujet de préoccupation et des mesures devraient être prises pour améliorer la concordance. De la même manière, les valeurs pour les autres paires peuvent être calculées comme suit : $\kappa(O1,O3) \approx 0.72$, $\kappa(O2,O3) \approx 0.22$. Les observateurs 1 et 3 affichent une bonne concordance. L'observateur 2 est clairement divergent et les raisons de cette différence nécessitent un examen plus approfondi (p. ex. envisager une formation supplémentaire).

8. Références

Cohen, J. (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20: 37-46.

Cohen, J. (1968) Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. Psychological Bulletin, 70(4): 213-220.

Bland, J. M. Altman D. G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement, Lancet: 307–310.

<http://www.seedtest.org/en/stats-tool-box-content--1--1143.html> (logiciel fondé sur la norme ISO 5725-2)

[Fin de l'annexe et du document]