
**Analyse sensorielle — Méthodologie —
Lignes directrices pour le contrôle de la
performance d'un jury sensoriel quantitatif**

*Sensory analysis — Methodology — Guidelines for monitoring the
performance of a quantitative sensory panel*

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO 11132:2012](https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012)

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>



iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO 11132:2012

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>



DOCUMENT PROTÉGÉ PAR COPYRIGHT

© ISO 2012

Droits de reproduction réservés. Sauf prescription différente, aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'ISO à l'adresse ci-après ou du comité membre de l'ISO dans le pays du demandeur.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Publié en Suisse

Sommaire

Page

Avant-propos	iv
1 Domaine d'application	1
2 Références normatives	1
3 Termes et définitions	1
4 Principe	3
5 Conditions expérimentales	4
6 Qualification des sujets	4
7 Mode opératoire	4
7.1 Contrôle par validation formelle de la performance	4
7.2 Analyse statistique des données par validation formelle de la performance (une seule session)	5
7.3 Performance globale du jury par validation formelle de la performance	7
7.4 Performance de chaque sujet par validation formelle de la performance	8
7.5 Problèmes de performance	10
7.6 Contrôle par des profils produits de routine	10
7.7 Plan expérimental pour l'étude de performance au fil du temps	10
7.8 Analyse statistique des données au fil du temps	10
7.9 Reproductibilité entre les jurys	11
7.10 Analyse statistique de profils complets	11
Annexe A (informative) Exemple d'application pratique	12
Annexe B (informative) Exemple d'utilisation de l'analyse CUSUM	19
Annexe C (informative) Exemple d'utilisation de la carte de Shewhart	22
Bibliographie	24

Avant-propos

L'ISO (Organisation internationale de normalisation) est une fédération mondiale d'organismes nationaux de normalisation (comités membres de l'ISO). L'élaboration des Normes internationales est en général confiée aux comités techniques de l'ISO. Chaque comité membre intéressé par une étude a le droit de faire partie du comité technique créé à cet effet. Les organisations internationales, gouvernementales et non gouvernementales, en liaison avec l'ISO participent également aux travaux. L'ISO collabore étroitement avec la Commission électrotechnique internationale (CEI) en ce qui concerne la normalisation électrotechnique.

Les Normes internationales sont rédigées conformément aux règles données dans les Directives ISO/CEI, Partie 2.

La tâche principale des comités techniques est d'élaborer les Normes internationales. Les projets de Normes internationales adoptés par les comités techniques sont soumis aux comités membres pour vote. Leur publication comme Normes internationales requiert l'approbation de 75 % au moins des comités membres votants.

L'attention est appelée sur le fait que certains des éléments du présent document peuvent faire l'objet de droits de propriété intellectuelle ou de droits analogues. L'ISO ne saurait être tenue pour responsable de ne pas avoir identifié de tels droits de propriété et averti de leur existence.

L'ISO 11132 a été élaborée par le comité technique ISO/TC 34, *Produits alimentaires*, sous-comité SC 12, *Analyse sensorielle*.

iTeh STANDARD PREVIEW (standards.iteh.ai)

ISO 11132:2012

<https://standards.iteh.ai/catalog/standards/sist/9c54bcd7-b6b9-4744-aea7-308079351781/iso-11132-2012>

Analyse sensorielle — Méthodologie — Lignes directrices pour le contrôle de la performance d'un jury sensoriel quantitatif

1 Domaine d'application

La présente Norme internationale donne des lignes directrices pour contrôler et évaluer la performance globale d'un jury descriptif quantitatif et la performance de chaque membre.

Un jury de sujets peut être utilisé comme un instrument pour évaluer l'intensité d'attributs sensoriels.

La performance est la mesure de la capacité d'un jury ou d'un sujet à effectuer des jugements valides d'attributs concernant les produits à évaluer. Elle peut être contrôlée à un moment donné ou suivie au fil du temps. La performance est la capacité d'un jury à détecter, identifier et mesurer un attribut, à utiliser les attributs de la même manière que d'autres jurys ou sujets, à différencier les stimulus, à utiliser correctement une échelle, à répéter ses propres résultats et à reproduire les résultats d'autres jurys ou sujets.

Les méthodes spécifiées permettent de contrôler et d'évaluer la cohérence, la répétabilité, l'absence de biais et la capacité de discrimination de jurys et de sujets. Le contrôle et l'évaluation de la concordance entre les membres du jury sont également abordés. Le contrôle et l'évaluation peuvent être réalisés au cours d'une session ou au fil du temps.

Le contrôle des données de performance permet au responsable du jury d'améliorer la performance du jury et des sujets, d'identifier les problèmes et les besoins de remise à niveau ou d'identifier les sujets qui ne sont pas suffisamment performants pour continuer à participer.

Les méthodes spécifiées dans la présente Norme internationale peuvent être utilisées par le responsable du jury pour évaluer de façon continue la performance du jury ou de chaque sujet.

La présente Norme internationale s'applique à des sujets ou à des jurys en formation aussi bien qu'à des jurys établis.

2 Références normatives

Les documents de référence suivants sont indispensables pour l'application du présent document. Pour les références datées, seule l'édition citée s'applique. Pour les références non datées, la dernière édition du document de référence s'applique (y compris les éventuels amendements).

ISO 5492, *Analyse sensorielle — Vocabulaire*

ISO 8586, *Analyse sensorielle — Lignes directrices générales pour la sélection, l'entraînement et le contrôle des sujets qualifiés et sujets experts*

ISO 8589, *Analyse sensorielle — Directives générales pour la conception de locaux destinés à l'analyse*

3 Termes et définitions

Pour les besoins du présent document, les termes et définitions donnés dans l'ISO 5492 ainsi que les suivants s'appliquent.

3.1

accord

capacité de jurys ou de sujets différents à attribuer des notes similaires à un attribut donné sur des échantillons du même produit

3.2
homogénéité
mesure de l'accord des réponses données entre différents sujets au cours d'une même session d'essai, entre sessions répétées pour un même jury de sujets ou entre sessions répétées pour un même sujet

3.3
biais du sujet
tendance d'un sujet à attribuer des notes qui, de manière constante, sont au-dessus ou en-dessous de la note vraie lorsque celle-ci est connue ou de la moyenne du jury lorsque celle-ci n'est pas connue

3.4
aberrant
jugement qui n'est pas conforme au modèle global des données ou qui est extrêmement différent des autres jugements donnés à des produits identiques ou similaires

3.5
dérive du jury
phénomène par lequel le jury, au fil du temps, change en sensibilité ou devient sensible à d'autres biais et qui, en conséquence, pour un produit de référence constant, modifie la position sur l'échelle sur laquelle un attribut est noté

3.6
performance
capacité d'un jury ou d'un sujet à réaliser des évaluations valides et fiables sur des stimulus et des attributs de stimulus

3.7
répétabilité
accord entre des évaluations effectuées sur des échantillons de produits équivalents, dans les mêmes conditions d'essai, par le même sujet ou le même jury

3.8
reproductibilité
accord entre des évaluations effectuées sur des échantillons de produits équivalents, dans des conditions d'essai différentes, avec différentes tâches ou par un sujet ou un jury différent

NOTE La reproductibilité peut être mesurée dans l'une des situations suivantes:

- la reproductibilité d'un jury à court terme, mesurée entre deux (ou plus) sessions séparées de plusieurs jours;
- la reproductibilité d'un jury à moyen ou long terme, mesurée entre des sessions séparées de plusieurs mois;
- la reproductibilité entre différents jurys, dans le même laboratoire ou dans des laboratoires différents;
- la reproductibilité d'évaluations effectuées par un seul sujet sur différents attributs d'un produit.

3.9
validation
processus consistant à établir que des données sensorielles sont corrélées avec d'autres données obtenues sur des échantillons du même produit (par exemple mesures de laboratoire, perception du consommateur, résultats d'autres jurys, plaintes du consommateur) ou qu'un jury ou un sujet est capable de remplir des critères de performance spécifiés

3.10
session
période pendant laquelle les produits sont évalués

NOTE Lors d'une même session, un ou plusieurs produits peuvent être évalués par un ou plusieurs sujets. Pour un sujet, qu'il soit seul ou qu'il fasse partie d'un jury, les sessions sont séparées dans le temps.

3.11
sessions répétées
sessions au cours desquelles les sujets, les produits, les conditions d'essai et la tâche sont identiques

4 Principe

La présente Norme internationale concerne les jurys sensoriels utilisés pour évaluer l'intensité d'un ou de plusieurs attributs sensoriels afin de réaliser des descriptions ou des profils quantitatifs de produits. Différentes méthodes sont appropriées pour évaluer et contrôler la performance des jurys utilisés pour analyser des différences.

La performance d'un jury sensoriel quantitatif peut être évaluée soit en utilisant les évaluations déjà disponibles soit à partir de sessions réalisées spécialement dans le but d'obtenir des données de performance.

La présente Norme internationale peut être utilisée soit pour le contrôle périodique, soit pour la vérification des données de profil en cours d'acquisition.

Une procédure de contrôle administrée à des intervalles périodiques est appropriée pour l'accréditation et pour d'autres fins. Un logigramme de cette procédure est donné à la Figure 1.

Pour examiner des données de profil générées par un jury, il peut être approprié d'utiliser des données qui proviennent d'expériences de profil bien différentes en employant différents types de produits, nombres de produits, etc. La procédure est la même que celle indiquée à la Figure 1. Cependant, comme il n'y a aucune différence prédéfinie, il est recommandé que les attributs qui sont significativement discriminés par l'ensemble du jury, pour un profil donné, soient ensuite utilisés comme mesures clés pour contrôler la performance de chaque membre du jury. Les attributs pour lesquels aucune différence n'est observée ne peuvent pas être utilisés de manière fiable pour contrôler la cohérence: en effet, l'absence de concordance intra et inter-membres du jury signifie que les produits sont très similaires pour ces caractéristiques.

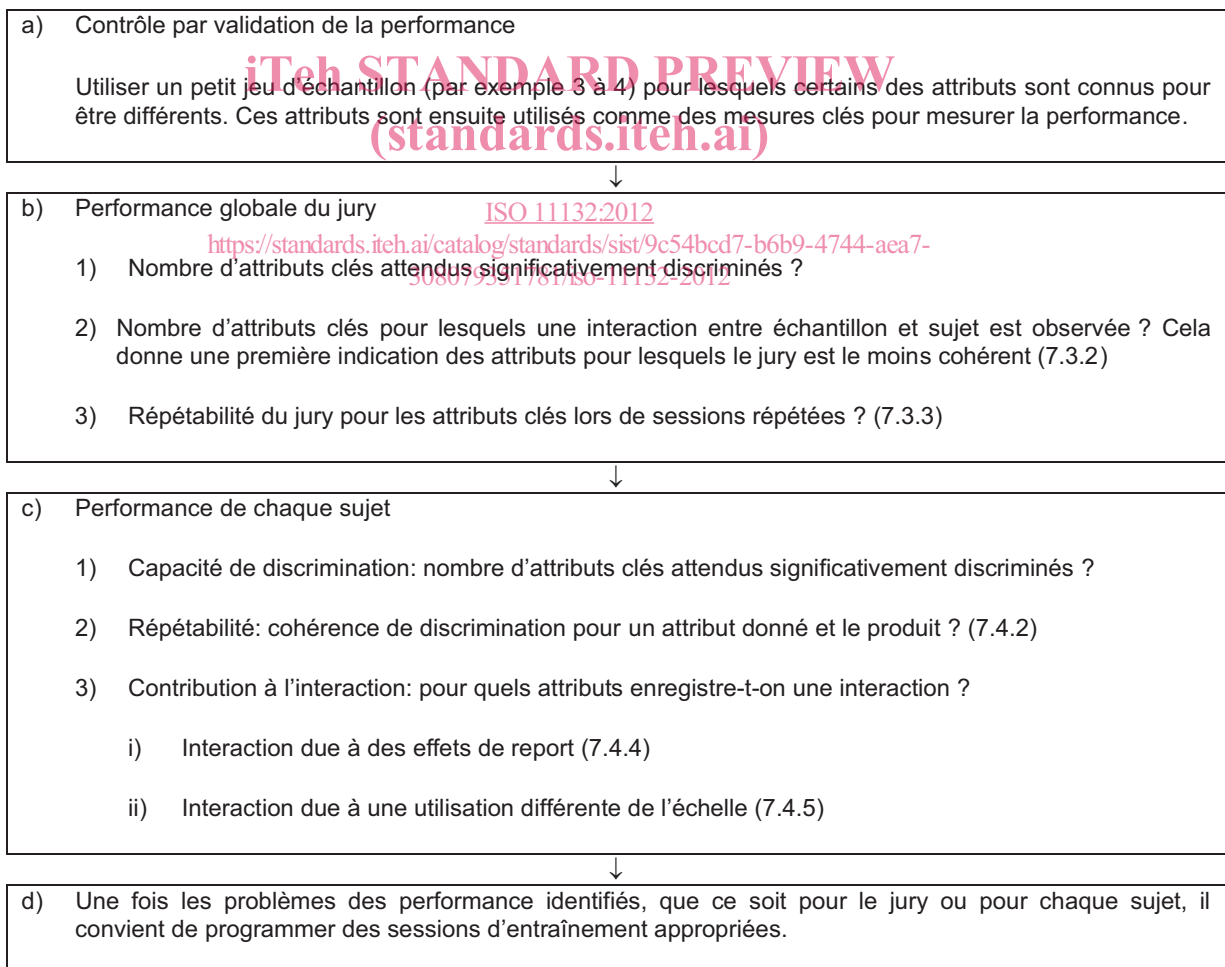


Figure 1 — Logigramme de contrôle de la performance

Les indicateurs qui peuvent être déterminés au cours de la même session sont les suivants.

- *Biais d'un sujet*, mesuré comme la différence entre la moyenne du sujet et une valeur «vraie» connue, ou la moyenne du jury utilisée comme une estimation de la valeur «vraie».
- *Répétabilité d'un sujet*, inversement reliée à l'écart-type calculé à partir d'évaluations répétées par le sujet sur le même échantillon ou à partir de répétitions effectuées sur le même produit.
- *Reproductibilité d'un sujet*, inversement reliée à l'écart-type des biais du sujet en fonction de chaque produit.
- *Discrimination d'un sujet*, mesurée comme la capacité à attribuer, de manière cohérente, des notes différentes à des produits différents.

Chez un sujet, le biais peut indiquer une acuité sensorielle et/ou une utilisation de l'échelle de réponses différente de celle des autres sujets.

S'il apparaît qu'un sujet réalise des évaluations différentes de celles des autres sujets, examiner l'ensemble des résultats afin de déterminer si:

- a) les évaluations sont cohérentes ou variables pour des échantillons répétés du même produit;
- b) les évaluations sont similaires ou différentes pour des échantillons de produits différents;
- c) le biais se produit avec la totalité, ou une partie seulement, des échelles d'évaluation.

L'analyse de la variance (ANOVA) peut être utilisée pour répondre à ces questions.

Dans certains cas, le biais peut indiquer un sujet de capacité supérieure dont les résultats sont particulièrement utiles. Dans d'autres cas, un sujet montrant un biais peut nécessiter un réentraînement ou il peut être exclu du jury.

Une seule approche cohérente pour l'analyse statistique des résultats est décrite ici. Toutefois, certains attributs de la performance du jury peuvent être évalués par plus d'une mesure descriptive. Par exemple, le carré moyen de l'erreur et l'erreur sur l'écart-type (sa racine carrée) expriment tous deux la variabilité dans l'évaluation d'un produit. Il convient que les mesures utilisées soient celles habituellement employées dans le domaine d'application.

D'autres mesures pertinentes concernant l'accord entre les sujets quant à l'utilisation de l'échelle pour un attribut sont l'interaction entre sujet et produit et le coefficient de corrélation entre les notes d'un sujet et les moyennes du jury. Il se peut qu'un sujet n'ait pas de biais mais qu'il utilise l'échelle de manière différente. Une corrélation proche de 1, une pente de régression proche de 1 et une ordonnée à l'origine de la droite de régression proche de 0 indiquent un bon accord entre un sujet et le reste du jury.

Avec un petit nombre de jugements (moins de six), il convient d'interpréter avec prudence le coefficient de corrélation car il peut être élevé (jusqu'à 0,7) par le seul fait du hasard.

5 Conditions expérimentales

Les installations d'essai doivent être conformes à l'ISO 8589.

6 Qualification des sujets

Le jury doit avoir au moins le niveau de qualification et d'expérience de sujets qualifiés (ISO 8586).

7 Mode opératoire

7.1 Contrôle par validation formelle de la performance

Lors de chaque session, il convient de présenter au jury de sujets un jeu d'échantillons similaires à ceux devant être examinés par le jury lors de l'évaluation des produits et pour lesquels des différences statistiquement significatives entre au moins deux échantillons peuvent être garanties pour au moins huit attributs.

Ce nombre est recommandé pour encourager les responsables du jury ou les responsables en analyse sensorielle à identifier et à sélectionner des échantillons de validation qui présentent une mesure réaliste et statistique de la performance d'un jury.

Ces attributs clés sont utilisés comme des mesures clés pour évaluer la performance du jury. Il convient que le jeu d'échantillons comprenne des répétitions. Le nombre de répétitions doit être identique pour chaque échantillon. Le nombre de sujets, d'échantillons et de répétitions dépend des produits, des attributs sensoriels évalués et de l'objectif du mode opératoire. Par exemple, 2 ou 3 répétitions de 3 ou 4 échantillons pourraient être utilisées. Il convient de veiller à limiter le nombre d'évaluations requises de façon à éviter la fatigue sensorielle. Il convient que les attributs des échantillons couvrent l'étendue des valeurs que le jury estime lors de l'évaluation des produits.

Un plan expérimental en blocs aléatoires a été adopté, les sujets étant les «blocs».

Si l'on attend un effet de report d'un échantillon à l'autre, un plan expérimental approprié est le carré latin de Williams. Le plan de base utilise quatre sujets et quatre échantillons.

Tableau 1 — Carré latin de Williams

Sujet	Ordre			
	1	2	3	4
1	A	B	C	D
2	B	D	A	C
3	C	A	D	B
4	D	C	B	A

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Dans ce plan, chaque sujet échantillonne les quatre produits dans un ordre différent et, pour chaque sujet, chaque produit particulier est suivi d'un produit différent. Par exemple, A est suivi de B pour le sujet 1, de C pour le sujet 2, de D pour le sujet 3 et d'aucun produit pour le sujet 4.

Si des multiples de quatre sujets sont disponibles, le même plan peut être répété pour chaque groupe de quatre.

7.2 Analyse statistique des données par validation formelle de la performance (une seule session)

Le Tableau 2 illustre une manière de disposer les résultats en tableaux et de les résumer. Certains logiciels informatiques peuvent nécessiter une organisation différente des données, par exemple avec les échantillons en colonnes et les sujets en lignes.

Tableau 2 — Résultats sujet par sujet

Échantillon	Sujet								Moyenne
	1		2		<i>j</i>		<i>n_q</i>		
	Notes	Moyenne	Notes	Moyenne	Notes	Moyenne	Notes	Moyenne	
1	Y_{111} Y_{112} Y_{11n_r}	$\bar{Y}_{11.}$			Y_{1j1} Y_{1j2} Y_{1jn_r}	$\bar{Y}_{1j.}$			$\bar{Y}_{1..}$
2									
<i>i</i>	Y_{i11} Y_{i12} Y_{in_r}	$\bar{Y}_{i1.}$			Y_{ij1} Y_{ij2} Y_{ijn_r}	$\bar{Y}_{ij.}$			$\bar{Y}_{i..}$
n_p									
Moyenne					$\bar{Y}_{.j.}$				$\bar{Y}_{...}$

Dans ce tableau, on suppose qu'il y a:

- n_p = nombre d'échantillons ($i = 1, 2 \dots n_p$);
- n_q = nombre de sujets ($j = 1, 2 \dots n_q$);
- n_r = nombre de répétitions par échantillon ($k = 1, 2 \dots n_r$).

ITeH STANDARD PREVIEW

Les mesures de la performance du jury dans son ensemble et de chacun des sujets, autres que le biais, demandent que les données soient analysées par une ANOVA.

Les détails des calculs de base ne sont pas indiqués dans le présente Norme internationale car les analyses sont normalement effectuées par un programme informatique.

Les données de chaque sujet sont analysées par une ANOVA à un facteur (Tableau 3).

Tableau 3 — ANOVA pour un sujet et un attribut

Source de variation	Degrés de liberté	Somme des carrés	Carré moyen	Rapport <i>F</i>
Entre les échantillons	$v_1 = n_p - 1$	S_1	$MS_1 = s_1/v_1$	$F = MS_1/MS_2$
Erreur	$v_2 = n_p(n_r - 1)$	S_2	$MS_2 = s_2/v_2$	
Total	$v_3 = n_p n_r - 1$	S_3		

n_p = nombre d'échantillons
 n_r = nombre de répétitions par échantillon

Les données pour la session complète sont analysées par une ANOVA en blocs aléatoires (Tableau 4).

Tableau 4 — ANOVA pour une session complète et un attribut

Source de variation	Degrés de liberté	Somme des carrés	Carré moyen	Rapport F
Entre les échantillons	$\nu_4 = n_p - 1$	S_4	$MS_4 = s_4 / \nu_4$	$F = MS_5 / MS_7^a$
Entre les sujets	$\nu_5 = n_q - 1$	S_5	$MS_5 = s_5 / \nu_5$	
Interaction	$\nu_6 = (n_p - 1)(n_q - 1)$	S_6	$MS_6 = s_6 / \nu_6$	$F = MS_6 / MS_7$
Erreur	$\nu_7 = n_p n_q (n_r - 1)$	S_7	$MS_7 = s_7 / \nu_7$	
Total	$\nu_8 = n_p n_q n_r - 1$	S_8		

n_p = nombre d'échantillons
 n_q = nombre de sujets
 n_r = nombre de répétitions par échantillon

^a Si l'interaction est significative, le rapport F pour la ligne «entre les sujets» est calculé par $F = MS_5 / MS_6$ avec le carré moyen de l'interaction au dénominateur.

7.3 Performance globale du jury par validation formelle de la performance

7.3.1 Discrimination des attributs clés

Il convient de déterminer la proportion des attributs clés qui ont été significativement discriminés comme attendu. Pour chaque attribut, cela est indiqué par la variation significative entre les échantillons pour un niveau de 0,05 dans le tableau ANOVA construit pour une session (Tableau 4). Plus la proportion d'attributs clés significativement discriminés est élevée, plus le jury est performant. Il convient que le jury soit ré-entraîné sur les attributs clés qui ne sont pas significativement discriminés comme attendu.

7.3.2 Homogénéité du jury

Un jury n'est pas homogène lorsque des sujets sont en désaccord avec le reste du jury.

Un jury n'est pas homogène si l'interaction entre échantillon et sujet dans l'ANOVA est significative pour un niveau de 0,05.

Le degré d'homogénéité du jury est inversement relié à l'écart-type de l'interaction, s_i .

$$s_i = \sqrt{\frac{MS_6 - MS_7}{n_r}}$$

Voir le Tableau 4.

Il convient de déterminer le nombre d'attributs clés donnant une interaction entre échantillon et sujet significative. Consulter le tableau ANOVA pour chaque attribut et noter ceux présentant une interaction pour un niveau de 0,05. Plus le nombre d'attributs clés donnant une interaction significative est élevé, moins le jury est homogène. Il convient que le jury soit ré-entraîné sur les attributs clés qui donnent une interaction significative.

7.3.3 Répétabilité du jury

La répétabilité du jury peut être estimée d'après la répétabilité de chaque sujet. Elle est inversement reliée à l'erreur sur l'écart-type, s_e :

$$s_e = \sqrt{MS_7}$$

Voir le Tableau 4.

7.3.4 Reproductibilité du jury

Pour vérifier la reproductibilité du jury, effectuer des évaluations d'autres échantillons de produits identiques au cours de sessions différentes.

Dans une ANOVA à trois facteurs (échantillons, sujets, sessions), il convient que la variation de la ligne «entre les sessions» ne soit pas significative pour un niveau de 0,05.

Il convient que l'interaction entre échantillons et sessions ne soit pas significative pour un niveau de 0,05. Une significativité indiquerait que l'évaluation des différences entre les échantillons aurait varié d'une session à l'autre.

Il convient que l'interaction entre sujets et sessions ne soit pas significative pour un niveau de 0,05. Une significativité indiquerait que les biais de chaque sujet auraient varié d'une session à l'autre.

Si l'analyse est utilisée pour décrire la performance du jury dans son ensemble, alors les facteurs de l'ANOVA (sessions, échantillons et sujets) sont des facteurs aléatoires. Les écarts-types des composantes peuvent être combinés pour donner une mesure de la reproductibilité.

Écart-type de reproductibilité, s_R :

$$s_R = \sqrt{s_e^2 + s_s^2 + s_{\text{sess}}^2 + s_{s \times \text{sess}}^2 + s_{\text{prod} \times \text{sess}}^2}$$

où

e signifie erreur;

s signifie sujets;

sess signifie sessions;

prod signifie produits.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO 11132:2012

Les estimations du biais et de la variation peuvent être présentées sous forme de tableaux et/ou de graphiques. Les graphiques en fonction de la durée indiquent si des dérives, des sauts ou des problèmes occasionnels sont survenus.

Des exemples de telles présentations sont l'analyse CUSUM (voir l'Annexe B) et les cartes de contrôle de Shewhart (voir l'Annexe C).

7.4 Performance de chaque sujet par validation formelle de la performance

7.4.1 Capacité de discrimination d'un sujet

La capacité de discrimination est mesurée par la proportion d'attributs clés ayant été significativement discriminés comme attendu. Pour chaque attribut, cela est indiqué par une variation significative de la ligne «entre les échantillons» pour un niveau de 0,05 dans le tableau ANOVA (Tableau 3). Plus la proportion d'attributs clés significativement discriminés est élevée, plus le sujet est performant. Il convient que le sujet soit ré-entraîné sur les attributs clés qui ne sont pas significativement discriminés comme attendu.

7.4.2 Répétabilité d'un sujet

La répétabilité d'un sujet est inversement liée à l'erreur sur l'écart-type du sujet, s_e :

$$s_e = \sqrt{MS_2}$$

Voir le Tableau 3.