

Adaptation interculturelle du schéma de codage des 4 habitudes en français pour évaluer les compétences de communication des médecins

Alexandre Bellier, Philippe Chaffanjon, Edward Krupat, Patrice François, José Labarère

Publié: 16 avril 2020 • <https://doi.org/10.1371/journal.pone.0230672>

Abstrait

Contexte

Le schéma de codage des quatre habitudes (4-HCS) est un instrument normalisé conçu pour évaluer les compétences de communication des médecins du point de vue d'un évaluateur externe, sur la base de consultations vidéo.

Objectif

Réaliser l'adaptation interculturelle du 4-HCS en français et évaluer ses propriétés psychométriques.

Méthodes

Le 4-HCS a été adapté de manière interculturelle en effectuant des traductions avant et arrière avec des traducteurs indépendants, conformément aux directives internationales. Quatre évaluateurs ont évalué 200 consultations d'étudiants en médecine enregistrées sur vidéo avec des patients standardisés, en utilisant la version française du 4-HCS. Nous avons examiné la cohérence interne, la structure factorielle, la validité de construit et la fiabilité du 4-HCS.

Résultats

Le score global moyen du 4-HCS était de 76,44 (écart-type, 12,34), sans effet de plancher ou de plafond dans les sous-échelles. La durée médiane de l'évaluation était de 8 minutes (intervalle de 4 à 19). L'alpha de Cronbach était de 0,94 pour le 4-HCS global, allant de 0,72 à 0,88 à travers les sous-échelles. Dans l'analyse factorielle confirmatoire, les statistiques de qualité d'ajustement n'ont pas corroboré la structure hypothétique à 4 habitudes. L'analyse factorielle exploratoire a abouti à deux dimensions, avec la fusion de trois habitudes conceptuellement liées en une seule dimension et un chargement croisé substantiel pour 15 des 23 éléments. Les estimations médianes moyennes des coefficients de corrélation intra-classe d'accord absolu étaient de 0,74 (intervalle de 0,68 à 0,84) et de 0,85 (intervalle de 0,76 à 0,91) pour la fiabilité inter et intra-juges des sous-échelles d'habitude, respectivement.

Conclusion

La version française du 4-HCS démontre une cohérence interne satisfaisante mais nécessite l'utilisation de deux évaluateurs indépendants pour obtenir une fiabilité acceptable. La structure factorielle sous-jacente de la version américaine originale et les adaptations interculturelles du 4-HCS méritent une étude plus approfondie.

Citation: Bellier A, Chaffanjon P, Krupat E, François P, Labarère J (2020) Adaptation interculturelle du schéma de codage des 4 habitudes en français pour évaluer les compétences de communication des médecins. *PLoS ONE* 15 (4): e0230672. <https://doi.org/10.1371/journal.pone.0230672>

Rédacteur: Gian Mauro Manzoni, Université eCampus, ITALIE

Reçu: 4 septembre 2019; **Accepté:** 5 mars 2020; **Publié:** 16 avril 2020

Copyright: © 2020 Bellier et al. Il s'agit d'un article en libre accès distribué sous les termes de la [licence d'attribution Creative Commons](#), qui permet une utilisation, une distribution et une reproduction sans restriction sur n'importe quel support, à condition que l'auteur original et la source soient crédités.

Disponibilité des données: L'ensemble de données est disponible dans Harvard Dataverse (voir à <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/IZTLWR>).

Financement: Cette étude a été soutenue par des subventions sans restriction de la Fondation d'Entreprise MACSF (19.951 €, AB: <https://www.macsf-exerciceprofessionnel.fr/Publications-actions-et-mecenat/Fondation-MACSF>). Les bailleurs de fonds n'ont joué aucun rôle dans l'étude conception, collecte et analyse des données, décision de publication ou préparation du manuscrit.

Intérêts concurrents: les auteurs ont déclaré qu'il n'y avait pas d'intérêts concurrents

introduction

Les compétences de communication des médecins sont des éléments clés de consultations médicales efficaces [1] et comprennent les compétences essentielles des médecins qui sont les plus recherchées par les patients [2]. Les preuves se sont accumulées, soutenant la conclusion qu'une communication de haute qualité est liée à une meilleure satisfaction des patients [3],

une plus grande observance du traitement [4], de meilleurs résultats pour la santé [5] et une diminution du risque de réclamations pour faute professionnelle [6]. De nombreuses organisations ont donc mis en place des programmes de formation structurés et évalué régulièrement les capacités de communication des médecins [3 , 7].

Le schéma de codage des quatre habitudes (4-HCS) est un instrument normalisé conçu pour évaluer 23 compétences ou comportements de communication des médecins du point de vue d'un évaluateur externe, sur la base de consultations vidéo [8]. Le 4-HCS est basé sur le cadre conceptuel du «Four Habits Model», un programme de formation qui a été développé au sein de l'US Kaiser Permanente Health Maintenance Organization et mis en œuvre pour enseigner des compétences de communication efficaces à des milliers de cliniciens de cette organisation au cours des deux dernières décennies [9]. Le modèle des quatre habitudes fait référence aux tâches d'entretien médical de base qui sont organisées en quatre dimensions à des fins didactiques, à savoir, investir au début (six éléments), susciter le point de vue du patient (trois éléments), faire preuve d'empathie (quatre éléments) et investir dans la fin (dix éléments) [9].

La version américaine originale du 4-HCS a démontré une fiabilité inter-évaluateurs acceptable et des preuves de validité de construit malgré une cohérence interne modérée, avec des coefficients alpha de Cronbach allant de 0,51 à 0,81 pour les quatre habitudes, plus de 100 rencontres médecin-patient enregistrées sur vidéo [8]. Depuis son développement initial, le 4-HCS a été utilisé en dehors du système Kaiser Permanente [9 - 13]. Des adaptations interculturelles du 4-HCS ont été publiées dans différentes langues: norvégien [14], allemand [15] et portugais brésilien [16]. Étant donné le potentiel du 4-HCS pour évaluer les compétences de communication de base et mesurer l'efficacité d'un programme de formation visant à modifier les compétences de communication [9], il est nécessaire de disposer d'une version française pouvant être utilisée avec les étudiants en médecine pendant la 4e année, programme de communication basé sur les compétences. À notre connaissance, seules trois études ont examiné les propriétés psychométriques du 4-HCS et aucune n'a étudié la structure factorielle sous-jacente [8 , 14 , 15]. Bien que les résultats des études précédentes soient prometteurs, les auteurs ont recommandé d'examiner plus en détail la validité et la fiabilité du 4-HCS dans différents contextes et populations [8].

Dans la présente étude, nous avons cherché à réaliser l'adaptation interculturelle du 4-HCS en français et à évaluer les propriétés psychométriques de la version adaptée, en utilisant les données originales de consultations vidéo-enregistrées d'étudiants en médecine avec des patients standardisés.

matériaux et méthodes

Étudier le design

Une procédure en deux étapes a été utilisée. Dans un premier temps, le 4-HCS a été traduit et adapté interculturellement en français. Deuxièmement, les propriétés psychométriques de la version française ont été étudiées, en ce qui concerne la cohérence interne, la validité et la fiabilité. La présente étude a été menée au CHU Grenoble Alpes et à l'École de médecine de l'Université Grenoble Alpes, France.

Adaptation interculturelle du 4-HCS en français

La traduction et l'adaptation interculturelle de la version source du 4-HCS et de son livre de codes ont été effectuées par Mapi Language Services, conformément aux directives publiées [17]. Mapi Language Services est une société internationale spécialisée dans le domaine de la traduction et de l'adaptation interculturelle des mesures des résultats rapportés par les patients (www.mapigroup.com/services/language-services/).

Le but du processus d'adaptation interculturelle était d'obtenir une traduction française qui était conceptuellement équivalente à la version source américaine, culturellement pertinente pour le contexte français et facilement compréhensible par les personnes qui utiliseraient l'instrument. À cette fin, nous avons utilisé une méthodologie rigoureuse impliquant la contribution du développeur 4-HCS sur des questions conceptuelles et un processus d'examen centralisé coordonné par un consultant expérimenté dans le domaine. Ce processus comprenait une compréhension commune des concepts 4-HCS par tous les participants impliqués dans le projet, le contrôle de la qualité par les traducteurs et une discussion sur les décisions de traduction à chaque étape.

En pratique, le développeur 4-HCS (EK) a été contacté pour obtenir l'autorisation d'utiliser et de traduire l'instrument et de l'inviter à participer au projet. Deux traducteurs qualifiés de langue maternelle française ont traduit indépendamment la version source du 4-HCS en français. Une seule version a été obtenue à l'issue d'une réunion de réconciliation des deux traducteurs. Ensuite, cette version a été rétro-traduite en anglais par un troisième anglophone qualifié, qui n'a pas vu la version américaine originale. La traduction arrière a été revue pour l'équivalence sémantique et opérationnelle par rapport à la version source du 4-HCS. Nous avons suivi une approche universaliste de l'équivalence, évaluant l'équivalence conceptuelle, item, sémantique, opérationnelle, de mesure et fonctionnelle [18]. Après que la version résultante a été prétestée par deux évaluateurs à l'aide de 63 consultations d'étudiants en médecine enregistrées sur vidéo avec des patients standardisés, des ajustements mineurs ont été apportés pour obtenir la version finale. Le chercheur principal a vérifié les preuves de la version finale et a corrigé les éventuelles erreurs.

Score 4-HCS

Conformément à la version source du 4-HCS, chaque élément a été noté sur une échelle de Likert à 5 points, allant de 1 à 5, les scores plus élevés indiquant une meilleure performance. Le point médian (c'est-à-dire 3) et les deux points finaux (c'est-à-dire 1 et 5) ont été ancrés, avec des descriptions comportementales spécifiques [8]. Les évaluateurs ont été encouragés à utiliser les catégories de point médian et de point final, avec d'autres catégories (c.-à-d. 2 et 4) à utiliser uniquement s'ils pensaient que les compétences en communication se situaient entre les points d'ancrage [8]. Cette approche a permis d'utiliser pleinement l'échelle de Likert en 5 points [8].

Un score global des compétences en communication a été calculé en additionnant les notes pour les éléments individuels, allant de 23 (c.-à-d. Moins efficace) à 115 (c.-à-d. Plus efficace). Quatre sous-échelles ont également été calculées, chacune correspondant à une dimension clé des compétences en communication (c.-à-d. Investir au début [intervalle, 6–30], Obtenir le point de vue du patient [intervalle, 3–15], Faire preuve d'empathie [intervalle, 4–20] et Investissez à la fin [range, 10–50]).

Échelle d'évaluation des compétences en relation médecin-patient

L'échelle d'évaluation des compétences en relation médecin-patient a été élaborée en français et validée au Canada [19]. Cette échelle se compose de 15 items explorant deux dimensions des compétences interpersonnelles, à savoir «compréhension de l'expérience de la maladie du patient» (huit items) et «communication efficace et respectueuse» (sept items). Chaque élément

pourrait être évalué sur une échelle de Likert à 4 points, allant de 15 (c.-à-d. Moins efficace) à 60 (c.-à-d. Plus efficace).

Echantillon d'étude et collecte de données

L'échantillon de l'étude consistait en des consultations d'étudiants en médecine enregistrées sur vidéo avec des patients standardisés. Les 218 étudiants en médecine ont été invités à participer dans le cadre du programme de 4 ans en communication basé sur les compétences. Ils ont été attribués par ordre alphabétique consécutif en fonction de leur nom de famille aux sessions du 1^{er} ou 2^e semestre qui ont eu lieu respectivement en octobre 2017 et avril 2018.

Les patients standardisés étaient 20 acteurs recrutés au Département des Arts du Spectacle de l'Université Grenoble Alpes. Sept scénarios de consultation médicale normalisés ont été élaborés et combinés avec neuf types de personnalité ou traits de caractère. Les acteurs ont été formés sur chaque scénario de consultation médicale par deux cliniciens au cours de séances de 2 h. Ensuite, ils ont été formés par leur instructeur d'improvisation sur chaque type de personnalité ou trait de caractère. Les combinaisons de scénarios de consultation médicale avec le type de personnalité ou le trait de caractère ont été prétestées dans le cadre d'une étude pilote.

Toutes les consultations médicales avec des patients standardisés ont été enregistrées sur vidéo à l'aide d'un équipement vidéo professionnel. L'enregistrement vidéo des consultations a permis de surmonter certains des défis de l'observation directe [20]. L'enregistrement vidéo a enregistré avec précision tous les événements survenus pendant les consultations, permettant aux évaluateurs de vérifier leurs observations autant de fois que nécessaire [20]. Les enregistrements vidéo pourraient être évalués par différents observateurs sans que les consultations ne soient interrompues [21]. Enfin, l'enregistrement vidéo a permis de fournir aux étudiants en médecine une rétroaction sur leur propre performance [20 , 22].

Quatre évaluateurs ont été recrutés pour le projet, dont un professeur titulaire de médecine (PC) et un résident en médecine (AB), tous deux ayant de l'expérience dans l'enseignement des compétences en communication, et deux étudiants en médecine. Ils ont suivi une session de formation qui a consisté à coder indépendamment cinq consultations enregistrées sur vidéo avec le 4-HCS et à discuter ensuite de ces cotes. Après avoir terminé la formation, les évaluateurs ont évalué les consultations de manière indépendante, chaque consultation étant notée par au moins trois évaluateurs différents. De plus, deux des évaluateurs ont évalué chaque consultation d'étudiant en médecine enregistrée sur vidéo deux fois, avec les deux évaluations à 2 mois d'intervalle, afin de quantifier la fiabilité intra-évaluateur.

Les données relatives à l'adaptation interculturelle du 4-HCS et à l'échelle d'évaluation des compétences en matière de relations médecin-patient ont été saisies à l'aide d'un formulaire de rapport de cas électronique. L'achèvement de chaque élément était obligatoire, il n'y avait donc aucune valeur manquante.

Taille de l'échantillon

Un échantillon de 200 consultations enregistrées sur vidéo était nécessaire pour l'analyse factorielle confirmatoire, basée sur des simulations précédentes [23]. En supposant une estimation ponctuelle du coefficient alpha de Cronbach proche de 0,80, cette taille d'échantillon fournirait une précision de $\pm 0,07$ (c.-à-d., Intervalle de confiance [IC] à 95% allant de 0,73 à 0,87).

En supposant une estimation ponctuelle du coefficient de corrélation intra-évaluateur de 0,80, nous avons estimé qu'un échantillon de 117 consultations enregistrées par vidéo atteindrait 80% de puissance pour démontrer qu'il serait supérieur à 0,70, avec un niveau de signification bilatéral de 0,05 [24]. En supposant une estimation ponctuelle du coefficient de corrélation inter-juges de 0,80 avec quatre évaluateurs, nous avons estimé qu'un échantillon de 68 consultations enregistrées sur vidéo atteindrait 80% de puissance pour démontrer qu'il serait supérieur à 0,70, avec un niveau de signification bilatéral de 0,05 [24].

analyses statistiques

Statistiques descriptives.

Pour tenir compte de la variabilité inter-juges, nous avons calculé les scores moyens de 4-HCS global et de sous-échelle des élèves [25]. Les scores globaux et de sous-échelle du 4-HCS ont été rapportés comme des moyennes avec les écarts-types. Les nombres et les pourcentages de réponses aux points d'ancrage pour les items et les scores globaux et sous-échelles ont été examinés pour détecter les effets de plancher ou de plafond. Les effets plancher et plafond inférieurs à 15% pour les scores de sous-échelle ont été jugés acceptables [26].

La cohérence interne.

La cohérence interne a été évaluée par la corrélation moyenne inter-items, la corrélation item-reste (c'est-à-dire la corrélation entre un item et le score qui a été formé par tous les autres items de la sous-échelle) et le coefficient alpha de Cronbach [27]. Le critère de cohérence interne a été rempli pour une corrélation item-repos > 0,40 et l'alpha de Cronbach > 0,70 a été jugé satisfaisant [26].

Structure interne.

La structure interne de l'adaptation interculturelle du 4-HCS en français a été vérifiée à l'aide d'une analyse factorielle confirmatoire et exploratoire, suivant les lignes directrices actuelles [27]. Dans un premier temps, une modélisation par équation structurelle des quatre habitudes prédéfinies a été réalisée avec les 23 items attribués aux habitudes prévues pour déterminer si les données de consultation médicale enregistrées par vidéo correspondent à la structure interne de la version américaine originale du 4-HCS. Cette modélisation par équation structurelle correspond à un modèle externe représentant les relations entre les variables latentes (c'est-à-dire les quatre habitudes prédéfinies) et les variables manifestes (c'est-à-dire les 23 items associés). Diverses statistiques de qualité d'ajustement ont été obtenues, y compris l'indice d'ajustement comparatif (CFI) [28], la valeur résiduelle moyenne quadratique standardisée (SRMR) et l'erreur quadratique moyenne d'approximation (RMSEA) ainsi que son IC à 90% [29]. Une valeur CFI de 0,90 ou plus et une valeur SRMR inférieure à 0,08 ont été considérées comme indiquant un ajustement satisfaisant du modèle. Une limite inférieure de l'IC à 90% pour l'estimation RMSEA inférieure à 0,05 ne rejeterait pas l'hypothèse selon laquelle l'ajustement était proche. Une borne supérieure supérieure à 0,10 ne rejeterait pas l'hypothèse selon laquelle l'ajustement était médiocre.

Deuxièmement, une analyse factorielle exploratoire a été effectuée afin d'examiner d'éventuelles structures alternatives à la version américaine originale du 4-HCS. Une méthode de rotation orthogonale (Varimax) des facteurs avec des valeurs propres supérieures à 1,00 a été utilisée, en supposant qu'ils étaient indépendants [30]. Les chargements primaires sur des dimensions prévues

supérieures à 0,40 avec des chargements croisés inférieurs à 0,30 ont été jugés satisfaisants.

La validité de construction.

Nous avons évalué la validité du concept en comparant les valeurs de score global et de sous-échelle du 4-HCS entre les consultations médicales enregistrées sur vidéo du premier et du deuxième semestre. Nous avons émis l'hypothèse que les scores moyens étaient plus élevés pour les consultations médicales enregistrées au cours du deuxième semestre. En effet, les étudiants du deuxième semestre étaient supposés être plus expérimentés dans les tâches d'entretien médical de base et plus sensibilisés aux compétences de communication que leurs homologues évalués au cours du premier semestre. Nous avons également examiné la validité convergente entre le 4-HCS et l'échelle d'évaluation des compétences de la relation médecin-patient, en utilisant les coefficients de corrélation de Pearson.

Fiabilité.

L'évaluation de la fiabilité inter et intra-juges était conforme aux *lignes directrices pour la déclaration des études de fiabilité et d'accord* (GRAAS) [31]. La fiabilité de la version française des scores globaux et des sous-échelles 4-HCS a été quantifiée par le coefficient de corrélation intra-classe (ICC) [32]. L'ICC convient aux études de fiabilité avec des conceptions déséquilibrées impliquant différents ensembles de noteurs [33]. Des ICC à la fois absolus et cohérents ont été calculés. Dans le cadre de l'approche de l'accord absolu, les notes étaient considérées en accord absolu si les scores 4-HCS de tous les évaluateurs correspondaient exactement [32]. Dans le cadre de l'approche de la cohérence de l'accord, les cotes étaient considérées comme cohérentes si les scores 4-HCS de deux évaluateurs différaient de la même valeur constante pour toutes les consultations enregistrées sur vidéo. Cela impliquait que les évaluateurs attribuaient le même classement à toutes les consultations enregistrées sur vidéo [34]. Les ICC individuels et moyens ont été estimés, les ICC moyens étant calculés sur deux évaluateurs. Bien que l'accord mesuré entre les évaluations individuelles soit plus courant, l'utilisation de CCI moyens a été indiquée dans cette étude parce que le 4-HCS était destiné à être utilisé par des équipes d'évaluateurs pour évaluer les consultations médicales enregistrées sur vidéo. Une valeur ICC égale ou supérieure à 0,70 indique une fiabilité satisfaisante [14].

Le protocole de l'étude a été approuvé par le Comité d'Ethique du Centre d'Investigation Clinique de Clermont-Ferrand, France (IRB 5891). Tous les participants ont reçu des informations du chercheur principal sur l'objectif général de l'étude et les exigences de confidentialité et ils ont ensuite fourni un consentement éclairé écrit.

Résultats

Mapi Language Services a traduit l'échelle 4-HCS en janvier 2018 et a émis un certificat de validation de traduction le 19 février 2018. Ils ont scrupuleusement respecté le protocole établi selon les normes internationales, produisant une adaptation interculturelle de l'échelle 4-HCS en conforme aux attentes. La version française du 4-HCS est présentée dans l' [annexe S1](#) . La version complète comprenait une traduction du «livre de codes», un guide de notation détaillé pour chaque élément avec des exemples de comportements et des évaluations suggérées.

Sur 218 étudiants en médecine admissibles, 200 (92%) ont participé à l'étude. Au total, 200 consultations avec des patients standardisés ont été enregistrées sur vidéo, dont 115 et 85 au cours des premier et deuxième semestres, respectivement ([Fig 1](#)). La durée médiane était de 8 minutes (intervalle de 4 à 19 minutes) par consultation enregistrée sur vidéo. L'évaluation de la fiabilité inter-juges a porté sur 800 notes, tandis que 400 notes ont contribué à l'évaluation de la fiabilité intra-évaluateur ([Fig 1](#)).

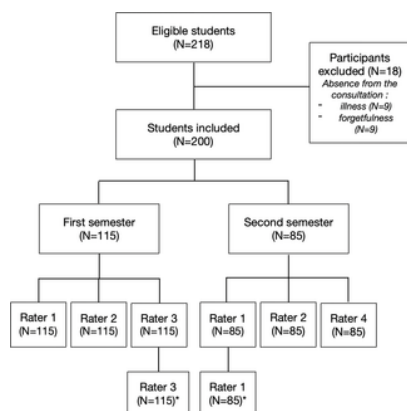


Fig 1. Flux d'étudiants en médecine et d'évaluateurs tout au long de l'étude.

<https://doi.org/10.1371/journal.pone.0230672.g001>

Le score moyen 4-HCS était de 76,44 (écart-type, 12,34) pour 200 consultations médicales enregistrées sur vidéo, sans effet de plancher ou de plafond observé pour les sous-échelles ([tableau 1](#)). Pourtant, les valeurs les plus élevées (5/5) et les plus basses (1/5) représentaient plus de 15% des répondants pour trois items et un item, respectivement. Les scores moyens 4-HCS variaient de 66,97 (ET, 10,26) sur la base de 200 consultations vidéo pour l'évaluateur 1 (c'est-à-dire, l'évaluateur le moins expérimenté) à 93,42 (ET 13,46) sur la base de 85 consultations enregistrées sur vidéo pour l'évaluateur 4 (c'est-à-dire, l'évaluateur le plus expérimenté).

Habit (range) Item	Mean score (SD)	Colling alpha (n)	Flow alpha (n)	Average inter item correlation	Item total correlation	Item rest correlation	Cronbach alpha ^a
1. Invest in the beginning [1-10]	18.77 (2.80)	0.99	0.99	0.48	0.99	0.99	0.99
2. Show flexibility	3.09 (0.93)	0.99	0.99	0.43	0.99	0.99	0.99
3. Invest wisely	3.98 (0.97)	0.99	0.99	0.48	0.99	0.99	0.99
4. Engage in small talk	2.66 (0.89)	0.99	0.99	0.46	0.99	0.99	0.99
5. Question role	2.25 (0.76)	0.99	0.99	0.52	0.99	0.99	0.99
6. Expressions of concern	3.12 (0.76)	0.99	0.99	0.50	0.99	0.99	0.99
7. Habit full agenda	3.07 (0.82)	0.99	0.99	0.47	0.99	0.99	0.99
8. Share the patient's perspective [1-11]	10.60 (2.04)	0.99	0.99	0.48	0.99	0.99	0.99
9. Patient's understanding	3.70 (0.76)	0.99	0.99	0.48	0.99	0.99	0.99
10. Goals for visit	2.81 (0.76)	0.99	0.99	0.50	0.99	0.99	0.99
11. Express on site	3.80 (0.96)	0.99	0.99	0.43	0.99	0.99	0.99
12. Demonstrate empathy [1-10]	11.75 (2.45)	0.99	0.99	0.52	0.99	0.99	0.99
13. Encourage emotional expression	3.33 (0.80)	0.99	0.99	0.49	0.99	0.99	0.99
14. Accept feelings	3.63 (0.75)	0.99	0.99	0.45	0.99	0.99	0.99
15. Identify feelings	2.80 (0.86)	0.99	0.99	0.50	0.99	0.99	0.99
16. Show good interpersonal behavior	4.19 (0.73)	0.99	0.99	0.53	0.99	0.99	0.99
17. Invest in the end [10-20]	10.91 (2.13)	0.99	0.99	0.46	0.99	0.99	0.99
18. Use patient's history of reference	3.60 (0.81)	0.99	0.99	0.50	0.99	0.99	0.99
19. Allow time to absorb	4.37 (0.76)	0.99	0.99	0.56	0.99	0.99	0.99
20. Give clear explanation	4.24 (0.76)	0.99	0.99	0.52	0.99	0.99	0.99
21. Offer options for tests	3.36 (0.80)	0.99	0.99	0.45	0.99	0.99	0.99
22. Test for comprehension	3.10 (0.86)	0.99	0.99	0.50	0.99	0.99	0.99
23. Summarize in decision	3.05 (0.82)	0.99	0.99	0.57	0.99	0.99	0.99
24. Engage plan acceptability	3.22 (0.86)	0.99	0.99	0.58	0.99	0.99	0.99
25. Engage barriers	2.02 (0.83)	0.99	0.99	0.46	0.99	0.99	0.99
26. Encourage questions	3.56 (0.81)	0.99	0.99	0.44	0.99	0.99	0.99
27. Plan for follow up	3.08 (0.81)	0.99	0.99	0.43	0.99	0.99	0.99
Overall (23-115)	76.80 (12.50)	0.99	0.99	0.47	0.99	0.99	0.99

Tableau 1. Statistiques récapitulatives et cohérence interne pour l'adaptation interculturelle du 4 Habit Coding Scheme en français (n = 200).
<https://doi.org/10.1371/journal.pone.0230672.t001>

L'alpha de Cronbach était de 0,94 (IC à 95%, 0,93 à 0,95) pour le 4-HCS, allant de 0,72 à 0,88 dans toutes les sous-échelles d'habitude ([tableau 1](#)). Tous les items sauf un remplissaient le critère de cohérence interne, avec des corrélations item-reste supérieures à 0,40. L'exception était l'élément 16 avec la corrélation élément-reste aussi faible que 0,15. La suppression de cet élément de la sous-échelle d'habitude «Investir à la fin» a amélioré l'alpha de Cronbach de 0,88 à 0,90.

Dans l'analyse factorielle confirmatoire, la modélisation par équation structurelle de 23 items répartis en quatre facteurs latents a donné des estimations CFI (0,79) et SRMR (0,09) qui n'ont pas atteint les seuils recommandés (CFI > 0,90 et SRMR < 0,08, respectivement). L'estimation RMSEA était de 0,12 (IC à 90%, 0,10 à 0,13), la borne inférieure de l'IC à 90% ne rejetant pas l'hypothèse que l'ajustement était proche et la borne supérieure de l'IC à 90% ne rejetant pas l'hypothèse que l'ajustement était médiocre.

L'analyse factorielle exploratoire des 23 items a identifié quatre composantes principales avec des valeurs propres supérieures à 1,0 et expliquant 66,8% de la variance globale ([tableau 2](#)). L'évaluation graphique du tracé d'éboulis a suggéré que l'instrument était proche de l'unidimensionnalité ([Annexe S2](#)), soutenant l'utilisation d'un score global 4-HCS. Pourtant, la structure factorielle de la version française du 4-HCS s'écartait de la structure à quatre dimensions hypothétique. Vingt et un éléments sur 23 présentaient des charges factorielles primaires supérieures à 0,40, tandis que 15 éléments produisaient des charges croisées supérieures à 0,30. L'analyse factorielle exploratoire de la version française du 4-HCS a abouti à la fusion de trois habitudes (à savoir, investir au début [les six éléments], susciter le point de vue du patient [deux sur trois éléments] et faire preuve d'empathie [les quatre items]) dans une seule dimension.

Habit Item	Factor 1	Factor 2	Factor 3	Factor 4
1. Invest in the beginning	.41			.40
2. Show flexibility	.48	.34		.35
3. Invest wisely	.58			
4. Engage in small talk	.38	.49	.32	
5. Question role	.74	.37		
6. Expressions of concern	.44	.32		
7. Habit full agenda				
8. Share the patient's perspective	.33	.48		.35
9. Patient's understanding	.48	.45	.47	.35
10. Goals for visit	.48			
11. Express on site				
12. Demonstrate empathy	.31			
13. Encourage emotional expression	.74	0.33		
14. Accept feelings	.46			
15. Identify feelings	.44	.37		
16. Show good interpersonal behavior				
17. Invest in the end	.33	.73		
18. Use patient's history of reference	.36			
19. Allow time to absorb				
20. Give clear explanation		.71		
21. Offer options for tests	.31	.48	.42	
22. Test for comprehension	.39	.36	.47	
23. Summarize in decision	.39	.79	.43	
24. Engage plan acceptability	.41	.41	.46	
25. Engage barriers	.48			
26. Encourage questions	.47	.47		
27. Plan for follow up	.36	.47		
Overall variance explained, %	46.5	8.1	8.2	6.8

Tableau 2. Analyse factorielle exploratoire pour les 23 items de l'adaptation interculturelle du Schéma de codage 4-Habit en français après rotation orthogonale Varimax (n = 200) .
<https://doi.org/10.1371/journal.pone.0230672.t002>

Selon l'hypothèse, tous les scores moyens d'habitude étaient significativement plus élevés pour les consultations d'étudiants en médecine enregistrées au deuxième semestre ([tableau 3](#)). La plupart des coefficients de corrélation de Pearson par paires entre les scores de l'échelle d'évaluation des compétences de l'habitude et de la relation médecin-patient étaient supérieurs à 0,70, indiquant une validité convergente satisfaisante ([tableau 4](#)).

Habit (range) Item	1 st semester (n = 115)	2 nd semester (n = 85)	p
1. Invest in the beginning [1-10]	17.14 (2.94)	15.81 (2.56)	< .001
2. Show the patient's perspective [1-11]	8.88 (1.91)	10.06 (2.02)	< .001
3. Demonstrate empathy [1-10]	11.60 (2.46)	11.96 (2.48)	< .001
4. Invest in the end [10-20]	11.81 (2.45)	11.81 (2.50)	< .001
Overall (23-115)	75.26 (12.48)	80.87 (16.77)	< .001

Tableau 3. Comparaison des scores du programme de codage 4-Habit pour les consultations d'étudiants en médecine enregistrées au premier et au deuxième semestre.
<https://doi.org/10.1371/journal.pone.0230672.t003>

Habit (range)	Physician-patient relationship competence assessment scale		
	1. Understanding of the patient's experience	2. Communication	3. Collaboration management
1. Score in the beginning	85	81	81
2. Effect the patient's perspective	79	81	79
3. Discussion empathy	84	84	84
4. Score in the end	79	79	79
Overall	80	79	80

Tableau 4. Corrélation de l'adaptation interculturelle du Schéma de codage 4-Habit en français et de l'échelle d'évaluation des compétences de la relation médecin-patient (n = 200) ¹.
<https://doi.org/10.1371/journal.pone.0230672.t004>

Aucune des quatre habitudes ne remplissait le critère de fiabilité inter-évaluateurs de 0,70, avec des estimations ponctuelles individuelles de l'ICC d'accord absolu allant de 0,42 à 0,64 ([tableau 5](#)). Il est intéressant de noter que le CCI d'accord absolu moyen pour le 4-HCS et trois habitudes sur quatre remplissaient le critère de fiabilité inter-juges de 0,70.

Habit	Inter-rater ICC (95% CI)			Intra-rater ICC (95% CI)		
	Individual	Average	Average	Individual	Average	Average
1. Score in the beginning	0,41	(0,30 to 0,52)	0,71	(0,63 to 0,79)	0,81	(0,72 to 0,90)
2. Effect the patient's perspective	0,42	(0,31 to 0,53)	0,68	(0,59 to 0,77)	0,81	(0,72 to 0,90)
3. Discussion empathy	0,61	(0,49 to 0,73)	0,71	(0,63 to 0,80)	0,81	(0,72 to 0,90)
4. Score in the end	0,64	(0,53 to 0,75)	0,81	(0,72 to 0,90)	0,71	(0,63 to 0,80)
Overall	0,61	(0,50 to 0,72)	0,71	(0,63 to 0,80)	0,81	(0,72 to 0,90)

Tableau 5. Estimations du coefficient de corrélation intra-classe d'accord absolu pour la fiabilité inter et intra-évaluateur pour l'adaptation interculturelle du schéma de codage 4-Habit en français (n = 200).
<https://doi.org/10.1371/journal.pone.0230672.t005>

L'ICC de concordance absolue individuelle était de 0,72 pour la fiabilité intra-évaluateur du score 4-HCS, avec des estimations ponctuelles allant de 0,45 à 0,71 selon les habitudes. L'ICC d'accord absolu moyen pour le 4-HCS et toutes les habitudes remplissait le critère de fiabilité intra-évaluateur de 0,70.

Après avoir limité notre échantillon analytique aux évaluateurs les plus expérimentés (évaluateurs n ° 1 et n ° 2), l'ICC d'accord absolu individuel était de 0,83 pour la fiabilité inter-juges et de 0,89 pour la fiabilité intra-évaluateur.

Discussion

L'utilisation d'instruments standardisés validés est préconisée pour évaluer les compétences interpersonnelles des médecins [8]. En effet, les instruments validés sont susceptibles de refléter avec précision le concept à mesurer, tandis que les instruments standardisés permettent des comparaisons à grande échelle des compétences de communication des médecins entre les études [30 , 35]. Le 4-HCS a donc été développé et validé à partir de plus de 1 025 consultations médicales enregistrées sur vidéo dans divers contextes aux États-Unis et en Europe occidentale ([tableau 6](#)). Pourtant, notre étude a été la première à examiner la structure factorielle sous-jacente du 4-HCS et à rendre compte de son adaptation interculturelle en français.

Author, year	Krupat, 2006	Ford-Innes, 2010	Chapin, 2011	Schell, 2014	Present study
Country	USA	Norway	USA	Germany	France
Setting	Hospital	Hospital	Family practice clinic	Primary and specialty clinics	School of medicine
Measurement period	2005-2006	2007-2008	2008-2009	2008-2009	2007-2008
Participants	Resident and senior medical students	Resident and senior medical students	Resident and senior medical students	Resident and senior medical students	Medical students
No. physicians	30	71	31	32	200
Standardized consultation	No	No	No	No	Yes
Recording	Video recorded	Video recorded	Video recorded	Audio taped	Video recorded
No. consultations	100	497	174	14	200
Inter-rater reliability					
1. Score in the beginning	0,77	-	0,81	0,81	0,84
2. Effect the patient's perspective	0,78	-	0,80	0,80	0,84
3. Discussion empathy	0,73	-	0,80	0,80	0,84
4. Score in the end	0,73	-	0,80	0,80	0,84
Overall	0,76	-	0,81	0,81	0,84
Physicians' factor analysis	Not performed	Not performed	Not performed	Not performed	4 principal components (60,0% of overall variance)
Construct validity					
1. Score in the beginning	0,71	-	-	0,81	0,84
2. Effect the patient's perspective	0,71	-	-	0,81	0,84
3. Discussion empathy	0,71	-	-	0,81	0,84
4. Score in the end	0,71	-	-	0,81	0,84
Overall	0,71	-	-	0,81	0,84
Intra-rater reliability					
1. Score in the beginning	0,79	-	0,80	0,80	0,84
2. Effect the patient's perspective	0,79	-	0,80	0,80	0,84
3. Discussion empathy	0,79	-	0,80	0,80	0,84
4. Score in the end	0,79	-	0,80	0,80	0,84
Overall	0,79	-	0,80	0,80	0,84
Inter-rater reliability					
1. Score in the beginning	0,83	-	-	0,87	0,89
2. Effect the patient's perspective	0,83	-	-	0,87	0,89
3. Discussion empathy	0,83	-	-	0,87	0,89
4. Score in the end	0,83	-	-	0,87	0,89
Overall	0,83	-	-	0,87	0,89

Tableau 6. Études primaires faisant état du développement ou de l'adaptation interculturelle du schéma de codage des 4 habitudes.
<https://doi.org/10.1371/journal.pone.0230672.t006>

Chacune des 200 consultations médicales enregistrées par vidéo dans cette étude a été notée sans données manquantes par trois évaluateurs différents sur quatre, reflétant l'acceptabilité et la faisabilité de la version française du 4-HCS. Bien que les consultations aient été simulées avec des patients standardisés et impliquaient des étudiants en médecine de 4 ans dans cette étude, les scores moyens globaux et de sous-échelle 4-HCS étaient cohérents avec les rapports précédents ([tableau 6](#)).

Étonnamment, la version française du 4-HCS donne de meilleures performances en termes de cohérence interne que l'instrument américain d'origine et les adaptations interculturelles précédentes. L'alpha de Cronbach était supérieur à 0,70 dans toutes les sous-échelles d'habitude (médiane 0,83; intervalle, 0,72–0,88) et se comparait favorablement à ceux rapportés dans le développement initial aux États-Unis (médiane, 0,66; intervalle, 0,51–0,81) et l'adaptation interculturelle allemande (médiane, 0,39; intervalle de 0,31 à 0,46) ([tableau 6](#)). Ce résultat contraste avec les études précédentes qui rapportent généralement de moins bonnes performances pour les adaptations interculturelles par rapport aux instruments standardisés originaux [30]. Krupat et coll. n'étaient pas préoccupés par la cohérence interne modérée de la version originale du 4-HCS et ont émis l'hypothèse qu'une formation réussie en communication se traduirait par une communication plus efficace et donc une plus grande cohérence interne [8]. Une autre explication potentielle de cet écart peut être que le cadre conceptuel du 4-HCS manque de généralisabilité ou de robustesse et ne s'applique donc pas de la même manière à toutes les expériences des populations cibles. Il est à noter que la version

française du 4-HCS a été utilisée pour évaluer les consultations d'étudiants en médecine avec des patients standardisés, tandis que la version américaine originale a été utilisée pour évaluer les consultations médicales des résidents et des cadres supérieurs ([tableau 6](#)).

À notre connaissance, aucune étude antérieure n'a examiné la structure factorielle du 4-HCS. Dans l'analyse factorielle confirmatoire, les statistiques de qualité d'ajustement ne soutenaient pas la structure à 4 dimensions hypothétique pour la version française du 4-HCS. L'analyse factorielle exploratoire a abouti à deux dimensions, avec la fusion de trois habitudes conceptuellement liées (investir au début, susciter le point de vue du patient, faire preuve d'empathie) dans une seule dimension. De plus, un chargement croisé substantiel a été observé pour 15 des 23 items, ce qui suggère que la structure factorielle sous-jacente de la version française du 4-HCS était discutable.

Ces résultats n'invalident pas nécessairement la structure postulée du 4-HCS. En effet, l'analyse factorielle ne peut discriminer que les constructions non corrélées dans un ensemble de données [36]. Le fait que trois habitudes soient corrélées entre elles dans l'échantillon de la présente étude n'implique pas que ces échelles mesurent le même concept [36]. Les étudiants en médecine participant à la présente étude n'ont pas été formés avec le modèle Four Habit avant que leurs compétences en communication ne soient évaluées à l'aide du 4-HCS. Cela pourrait expliquer pourquoi la structure factorielle de la version française s'écartait de la structure à quatre dimensions postulée du 4-HCS. Pourtant, nous ne pouvons pas exclure que ce problème soit inhérent au 4-HCS original plutôt que spécifique à notre échantillon d'étude. Aucune étude antérieure (y compris l'étude de développement originale) n'a examiné la structure factorielle du 4-HCS. Par conséquent, il est nécessaire de poursuivre l'examen de la structure factorielle sous-jacente de la version américaine d'origine du 4-HCS.

Un item (16. *Donnez une explication claire*) n'a été attribué à aucune dimension empirique dans l'analyse factorielle exploratoire et a également produit la corrélation item-repos la plus faible, détériorant la cohérence interne de l'échelle d'habitude correspondante. Au total, ces deux observations remettent en question la pertinence de ce point, et donc sa suppression de la version française du 4-HCS devrait être discutée.

La preuve de la validité convergente a été fournie par la corrélation attendue entre le 4-HCS et les scores de l'échelle d'évaluation des compétences de la relation médecin-patient. Notre observation selon laquelle les étudiants en médecine plus expérimentés dans les tâches d'entretien de base ont donné des scores plus élevés pour les quatre habitudes a soutenu la validité de construit du 4-HCS.

Les comparaisons des estimations d'inter-fiabilité et d'intra-fiabilité ont été confondues par l'hétérogénéité inter-études dans les types de coefficients de corrélation utilisés. Seules deux études, dont la présente, ont utilisé un CCI d'accord absolu ([tableau 6](#)). L'ICC est une alternative recommandée à la corrélation du coefficient r de Pearson pour évaluer l'inter ou l'intra-fiabilité [14 , 31]. L'ICC médian d'accord absolu pour la fiabilité inter-juges (0,49, intervalle, 0,42–0,64) et intra-évaluateur (0,58, intervalle, 0,45–0,71) pour la version française du 4-HCS était inférieur à ceux rapportés en allemand étude d'adaptation interculturelle (0,81, intervalle, 0,78–0,85 et 0,83, intervalle, 0,72–0,87, respectivement). Une fiabilité inter-évaluateurs plus faible pourrait être expliquée par l'utilisation de consultations simulées avec des patients standardisés, l'utilisation de consultations vidéo plutôt que sur bande audio, ou des niveaux variables d'expérience des évaluateurs dans notre étude. Nous avons observé que plus le niveau d'expérience de l'évaluateur est élevé, plus les scores 4-HCS sont élevés dans notre étude. Des associations incohérentes ont été signalées entre les scores d'expérience ou d'ancienneté et de communication dans les examens cliniques objectifs structurés [37]. Après avoir limité notre échantillon analytique aux évaluateurs les plus expérimentés (c.-à-d. Les évaluateurs n° 1 et n° 2), les estimations de l'ICC étaient supérieures à 0,80 pour la fiabilité inter-juges.

Fait intéressant, une amélioration substantielle de la fiabilité inter-juges (médiane, 0,74; intervalle, 0,68–0,84) et intra-évaluateur (0,85, intervalle, 0,76–0,91) des scores d'habitudes a été obtenue en utilisant les estimations CCI moyennes en accord absolu ([tableau 5](#)). Cette dernière constatation appuie la nécessité d'utiliser deux évaluateurs indépendants pour évaluer les compétences de communication avec le 4-HCS sur la base de consultations médicales enregistrées sur vidéo, dans la pratique de routine.

Cette étude a des implications potentielles pour l'évaluation de routine des compétences de communication des médecins à l'aide de l'adaptation interculturelle du 4-HCS en français. Premièrement, cette étude fournit des preuves sur la validité de l'échelle 4-HCS pour des consultations simulées avec des patients standardisés. Deuxièmement, la version française de l'échelle 4-HCS a démontré une cohérence interne, qui était encore plus élevée que la version américaine d'origine, permettant des comparaisons internationales. Troisièmement, notre étude a remis en question la structure factorielle sous-jacente hypothétique du 4-HCS. Étant donné que le 4-HCS a été développé à l'origine aux États-Unis, le cadre conceptuel et la structure des facteurs peuvent manquer de généralisabilité ou de robustesse et peuvent donc ne pas s'appliquer également à d'autres pays. L'étude de la structure factorielle de la version américaine d'origine est nécessaire pour résoudre ce problème. Quatrième,

Cette étude sur l'adaptation interculturelle comporte quelques mises en garde qui doivent être prises en compte. Premièrement, de vraies rencontres avec des patients auraient été préférables à des rencontres standardisées avec des patients pour évaluer les capacités de communication, en raison de leur authenticité [38]. En effet, les consultations simulées avec des patients standardisés diffèrent des consultations réelles de patients à bien des égards [39]. Les patients simulés ne souffrent pas de maladie et essaient seulement de le représenter à travers leur jeu. De plus, le recrutement et la formation de patients standardisés prennent du temps afin de produire une simulation de haute qualité [40]. Bien que notre étude n'ait pas exploré de vraies rencontres de patients, les patients standardisés nous ont permis de fournir à un grand nombre d'étudiants des scénarios cliniques reproductibles et cohérents du même niveau de difficulté [41].

Deuxièmement, notre étude a été menée auprès d'étudiants en médecine de 4^e année dans un seul hôpital affilié à une université et les résultats peuvent ne pas s'appliquer à d'autres contextes. Un éventail plus large de participants renforcerait la confiance dans les propriétés psychométriques de cette adaptation interculturelle du 4-HCS en français.

Troisièmement, le 4-HCS et l'échelle d'évaluation des compétences de la relation médecin-patient ont été remplis par le même évaluateur, de sorte qu'un effet de halo ne peut être exclu lors de l'évaluation de la validité convergente des deux instruments. Démontrer que les scores 4-HCS sont en corrélation avec l'expérience (standardisée) rapportée par les patients des compétences de communication du médecin fournirait une preuve plus solide de la validité du concept.

Quatrièmement, l'intervalle d'un semestre séparant les deux groupes d'étudiants pourrait être trop court pour évaluer la relation entre les scores 4-HCS et l'expérience en communication et les tâches d'entretien de base, bien que les associations soient significatives. La capacité du 4-HCS à discriminer les sujets avec différents niveaux de compétences en communication justifie une enquête plus approfondie. La sensibilité de la version française du 4-HCS aux évolutions des capacités de communication reste également à être documentée par des études longitudinales.

Conclusions

La version française du 4-HCS démontre une cohérence interne satisfaisante mais une fiabilité modérée, nécessitant l'utilisation de deux évaluateurs indépendants pour évaluer les capacités de communication des étudiants en médecine sur la base de consultations vidéo avec des patients standardisés. La structure factorielle empirique de la version française n'est pas conforme aux habitudes hypothétiques du 4-HCS original. Que cela reflète un problème spécifique avec notre échantillon d'étude d'adaptation interculturelle ou un problème plus général avec l'instrument n'est pas clair et mérite une enquête plus approfondie.

Renseignements à l'appui

Annexe S1. Adaptation interculturelle de l'échelle 4-HCS en français.

<https://doi.org/10.1371/journal.pone.0230672.s001>

(DOCX)

Annexe S2. Scree plot des 23 items de l'échelle 4-HCS.

<https://doi.org/10.1371/journal.pone.0230672.s002>

(TIF)

Remerciements

Les auteurs remercient Pauline Bouchet et Ariane Martinez, Université Grenoble Alpes, pour leur collaboration à la mise en place de consultations médicales simulées avec des patients standardisés, Noémie Kaladzé et Alexis Dechosal, Université Grenoble Alpes, pour la notation des consultations médicales enregistrées sur vidéo. Les auteurs remercient Alison Foote, du CHU Grenoble Alpes, et Linda Northrup pour leur aide dans l'édition du manuscrit en anglais.

Les références

1. Ong LML, de Haes JCJM, Hoos AM, Lammes FB. Communication médecin-patient: une revue de la littérature. *Soc Sci Med*. 1995; 40: 903–918. pmid: 7792630
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
2. Simpson SA, Shaw C, McNamara R. Quelle est la manière la plus efficace de maintenir la perte de poids chez les adultes? *BMJ*. 2011; 343: d8042. pmid: 22205707
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
3. Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. La formation en techniques de communication pour les médecins améliore la satisfaction des patients. *J Gen Intern Med*. 2016; 31: 755–761. pmid: 26921153
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
4. Zolnierek KBH, Dimatteo MR. Communication du médecin et adhésion du patient au traitement: une méta-analyse. *Soins médicaux*. 2009; 47: 826–834. pmid: 19584762
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
5. Kelley JM, Kraft-Todd G, Schapira L, Kossowsky J, Riess H. Influence de la relation patient-clinicien sur les résultats des soins de santé: une revue systématique et une méta-analyse d'essais contrôlés randomisés. *PloS One*. 2014; 9: e94207. pmid: 24718585
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
6. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Communication médecin-patient. La relation avec les réclamations pour faute professionnelle chez les médecins de soins primaires et les chirurgiens. *JAMA*. 1997; 277: 553–559. pmid: 9032162
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
7. Brown JB, Boles M, Mullooly JP, Levinson W. Effet de la formation des cliniciens à la communication sur la satisfaction des patients. Un essai contrôlé randomisé. *Ann Intern Med*. 1999; 131: 822–829. pmid: 10610626
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
8. Krupat E, Frankel R, Stein T, Irish J. The Four Habits Coding Scheme: validation d'un instrument pour évaluer le comportement de communication des cliniciens. *Conseillers en éducation des patients* 2006; 62: 38–45. pmid: 15964736
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
9. Stein T, Frankel RM, Krupat E. Améliorer les compétences de communication des cliniciens dans une grande organisation de soins de santé: une étude de cas longitudinale. *Conseillers en éducation des patients* 2005; 58: 4–12. pmid: 15950831
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
- dix. Fossli Jensen B, Gulbrandsen P, Dahl FA, Krupat E, Frankel RM, Finset A. Efficacité d'un cours de courte durée en communication clinique pour les médecins hospitaliers: résultats d'un essai contrôlé randomisé croisé (ISRCTN22153332). *Conseillers en éducation des patients* 2011; 84: 163–169. pmid: 21050695
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
11. Blanch-Hartigan D, Hall JA, Roter DL, Frankel RM. Biais sexiste dans les perceptions des patients des comportements centrés sur le patient. *Conseillers en éducation des patients* 2010; 80: 315–320. pmid: 20638813
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
12. Gulbrandsen P, Madsen HB, Benth JS, Laerum E. Les prestataires de soins de santé communiquent moins bien avec les patients souffrant de lombalgie chronique - une étude de rencontres dans une clinique de mal de dos au Danemark. *Douleur*. 2010; 150: 458–461. pmid: 20705213
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)

13. Clayton MF, Latimer S, Dunn TW, Haas L. Évaluation de la communication centrée sur le patient dans un contexte de médecine familiale: comment la mesurer et quelle opinion compte? *Conseillers en éducation des patients* 2011; 84: 294-302. pmid: 21733653
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
14. Fossli Jensen B, Gulbrandsen P, Benth JS, Dahl FA, Krupat E, Finset A. Fiabilité inter-évaluateur pour le schéma de codage des quatre habitudes dans le cadre d'un essai contrôlé randomisé. *Conseillers en éducation des patients* 2010; 80: 405-409. pmid: 20685066
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
15. Scholl I, Nicolai J, Pahlke S, Kriston L, Krupat E, Härter M. La version allemande du Four Habits Coding Scheme - association entre la communication des médecins et le partage des compétences décisionnelles lors de la rencontre médicale. *Conseillers en éducation des patients* 2014; 94: 224-229. pmid: 24286733
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
16. Catani RR, Valadares EDS, Lacombe JB, Mendonça TM da S, Silva CHM da, Paro HBM da S. *Cad Saude Publica*. 2018; 34: e00013918. pmid: 30427408
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
17. Epstein J, Santo RM, Guillemin F. Un examen des lignes directrices pour l'adaptation interculturelle des questionnaires n'a pas pu aboutir à un consensus. *J Clin Epidemiol*. 2015; 68: 435-441. pmid: 25698408
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
18. Harachi TW, Choi Y, Abbott RD, Catalano RF, Bliesner SL. Examen de l'équivalence des concepts et des mesures dans divers échantillons. *Prev Sci*. 2006; 7: 359-368. pmid: 16845592
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
19. Côté L, Savard A, Bertrand R. [Évaluation de la compétence relation médecin-patient. Développement et validation d'un instrument d'évaluation]. *Can Fam Physician*. 2001; 47: 512-518. pmid: 11281084
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
20. Asan O, Montague E. Utilisation de méthodes de recherche par observation vidéo dans les rencontres de soins de santé primaires pour évaluer des interactions complexes. *Informez Prim Care*. 2014; 21: 161-170. pmid: 25479346
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
21. Ram P, Grol R, Rethans JJ, Schouten B, van der Vleuten C, Kester A. Évaluation des médecins généralistes par observation vidéo des performances communicatives et médicales dans la pratique quotidienne: questions de validité, de fiabilité et de faisabilité. *Med Educ*. 1999; 33: 447-454. pmid: 10354322
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
22. Nilsen S, Baerheim A. Commentaires sur les consultations vidéo enregistrées dans l'enseignement de la médecine: pourquoi les étudiants le détestent et l'aiment - une étude qualitative basée sur des groupes de discussion. *BMC Med Educ*. 2005; 5: 28. pmid: 16029509
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
23. Rouquette A, Falissard B. Exigences relatives à la taille des échantillons pour la validation interne des échelles psychiatriques. *Int J Methods Psychiatr Res*. 2011; 20: 235-249. pmid: 22020761
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
24. Walter SD, Eliasziw M, Donner A. Taille de l'échantillon et conceptions optimales pour les études de fiabilité. *Stat Med*. 17: 101-110. pmid: 9463853
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
25. Bland JM, DG Altman. Corrélation, régression et données répétées. *BMJ*. 1994; 308: 896. pmid: 8173371
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
26. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Des critères de qualité ont été proposés pour les propriétés de mesure des questionnaires sur l'état de santé. *J Clin Epidemiol*. 2007; 60: 34-42. pmid: 17161752
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
27. Nunnally JC, Bernstein I. *Théorie psychométrique*. 3e éd. New York: McGraw-Hill; 1994.
28. Bentler PM. Index d'ajustement comparatifs dans les modèles structurels. *Psychol Bull*. 1990; 107: 238-246. pmid: 2320703
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
29. Hancock GR, Mueller RO. *Modélisation d'équations structurelles: un deuxième cours*. CT: l'ère de l'information. Greenwich; 2006.
30. Boussat B, François P, Gandon G, Giai J, Seigneurin A, Perneger T, et al. Incohérences entre deux adaptations interculturelles de l'enquête hospitalière sur la culture de la sécurité des patients en français. *J Patient Saf*. 2017. pmid: 29140887
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
31. Kottner J, Audigé L, Brorson S, Donner A, Gajewski BJ, Hróbjartsson A, et al. Des lignes directrices pour les études de fiabilité et d'accord (GRRAS) ont été proposées. *J Clin Epidemiol*. 2011; 64: 96-106. pmid: 21130355
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)

32. Shrout PE, Fleiss JL. Corrélations intraclassées: utilisations pour évaluer la fiabilité de l'évaluateur. *Psychol Bull.* 1979; 86: 420–428. pmid: 18839484
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
33. Hallgren KA. Calcul de la fiabilité inter-évaluateurs pour les données d'observation: un aperçu et un didacticiel. *Tutor Quant Methods Psychol.* 2012; 8: 23–34. pmid: 22833776
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
34. McGraw KO, P S. Formant des inférences sur certains coefficients de corrélation intraclassée. *Méthodes Psychol.* 1996; 1: 30–46.
[Voir l'article](#) • [Google Scholar](#)
35. Gjersing L, Caplehorn JRM, Clausen T. Adaptation interculturelle des instruments de recherche: langage, contexte, temps et considérations statistiques. *BMC Med Res Methodol.* 2010; 10: 13. pmid: 20144247
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
36. Perneger TV, Leplège A, Etter JF. Adaptation interculturelle d'un instrument psychométrique: deux méthodes comparées. *J Clin Epidemiol.* 1999; 52: 1037-1046. pmid: 10526997
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
37. Chong L, Taylor S, Haywood M, Adelstein BA, Shulruf B. L'ancienneté et l'expérience des examinateurs sont associées à des préjugés lors de la notation des compétences en communication, mais pas aux compétences d'examen dans les examens cliniques objectifs structurés (ECOS) en Australie. *J Educ Eval Health Prof.* 2018. pmid: 30016854
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
38. Bokken L, Rethans JJ, Jöbsis Q, Duvivier R, Scherpbier A, van der Vleuten C. Instructivité de vrais patients et de patients simulés dans la formation médicale de premier cycle: une expérience randomisée. *Acad Med J Assoc Am Med Coll.* 2010; 85: 148-154. pmid: 20042841
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
39. Bokken L, Rethans JJ, Scherpbier AJJA, van der Vleuten CPM. Forces et faiblesses des patients simulés et réels dans l'enseignement des compétences aux étudiants en médecine: un bilan. *Simul Healthc J Soc Simul Healthc.* 2008; 3: 161-169. pmid: 19088660
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
40. Barrows HS. Un aperçu des utilisations des patients standardisés pour l'enseignement et l'évaluation des compétences cliniques. *AAMC. Acad Med J Assoc Am Med Coll.* 1993; 68: 443–453. pmid: 8507309
[Voir l'article](#) • [PubMed / NCBI](#) • [Google Scholar](#)
41. Collins RMH JP AMEE Medical Education Guide No. 13: vrais patients, simulateurs de patients et simulateurs lors d'examens cliniques. *Med Teach.* 1998; 20: 508-521.
[Voir l'article](#) • [Google Scholar](#)