

L'évaluation formative de la compétence interculturelle: Regard sur le développement d'un outil alternatif

Lilia A. Simões Forte, lfort089@uottawa.ca, Université d'Ottawa

Eric Dionne, Ph.D, eric.dionne@uottawa.ca, Université d'Ottawa

Résumé: Dans cet article, nous portons un regard analytique et critique sur les outils utilisés pour le développement et l'évaluation de la compétence interculturelle des étudiants universitaires. Plus particulièrement, nous présentons : (1) les avantages ainsi que les limites des instruments d'auto-évaluation fréquemment employés pour évaluer une telle compétence, (2) un outil alternatif visant à évaluer le jugement des étudiants dans une situation complexe du point de vue culturel et (3) une réflexion concernant l'utilité de cet instrument dans les pratiques d'évaluation formative.

Mots-clés : instruments, compétence interculturelle, validité, évaluation formative

1 Introduction

L'arrivée d'étudiants étrangers dans les universités pose divers défis pour l'enseignement et pour l'apprentissage, dans la mesure où les salles de classe deviennent de plus en plus hétérogènes du point de vue culturel (De Vita, 2007). En conséquence, de nombreux auteurs s'intéressent au développement et à l'évaluation de la compétence interculturelle (CI) (Byram, 1997; Deardorff, 2006; Franklin-Craft, 2010; Meade, 2010). Cette dernière peut être développée dans une variété de contextes éducatifs (formels ou informels) et se caractérise par ses composantes cognitives, affectives et comportementales. Elle suppose la mobilisation de ces composantes afin d'établir une interaction interpersonnelle positive avec des personnes d'autres cultures.

Une grande variété d'outils peut être utilisée pour évaluer le développement de cette compétence. À titre d'exemple, les formateurs¹ peuvent administrer des instruments d'auto-évaluation (exemples : questionnaires, échelles, sondages, etc.), des exercices individuels ou de groupe, des assimilateurs de culture ou encore procéder à des simulations, des observations ou à des entretiens (voir Fantini, 2009; Fowler et Mumford, 1999a, 1999b; Kohls et Knight, 1994; Paige, 2004; Tomalin, 2009). Il semble évident qu'il existe de nombreux choix quant à la nature des instruments d'évaluation d'une telle compétence. Par ailleurs, ils sont administrés dans une perspective d'évaluation diagnostique, formative ou sommative, en fonction du contexte ou des objectifs de formation. Cependant, des lacunes ont été identifiées dans le processus de validation de nombreux instruments. Aussi, peu d'auteurs se sont intéressés à la qualité du processus ou encore aux impacts de ce processus au regard de la fonction (diagnostique, formative, sommative) de l'évaluation (voir Mendenhall *et al.*, 2004; Paige, 2004).

Ce texte vise à examiner la portée et les limites des instruments fréquemment utilisés pour évaluer la CI, ainsi que de présenter l'exemple d'un outil alternatif d'évaluation d'une telle compétence. Pour concrétiser ces objectifs, cet article est divisé en quatre parties, notamment : (1) définition de la compétence interculturelle, (2) description des menaces à la validité dans l'évaluation de cette compétence, (3) présentation d'un outil visant à évaluer le jugement des étudiants dans une situation

¹ Le masculin est utilisé sans discrimination dans le but d'alléger le texte.

complexe du point de vue culturel, et (4) réflexion sur l'utilité de cet instrument dans un contexte d'évaluation formative.

2 Définition de la compétence interculturelle

Spitzberg et Changnon (2009) présentent près d'une quarantaine de modèles contemporains explicatifs de la CI. Ils peuvent être classifiés en cinq catégories. Premièrement, certains modèles sont uniquement composés de dimensions (ex. affective, comportementale, cognitive) ou de concepts et n'impliquent pas de liens entre ceux-ci (exemple : modèle en pyramide des compétences interculturelles de Deardorff, 2006). Deuxièmement, les modèles co-orientationnels établissent des relations entre les concepts afin de mieux comprendre le processus et les résultats qui en découlent (exemple : modèle de Byram, 1997). Troisièmement, les modèles développementaux considèrent que la CI évolue au long du temps, individuellement ou collectivement (exemple: modèle de développement des compétences interculturelles de Bennett, 1986). Ensuite, les modèles adaptatifs accordent une attention particulière au processus d'adaptation (exemple : modèle de la communication interculturelle de Kim, 2001). Finalement, les modèles de causalité considèrent la CI comme un système linéaire (exemple : modèle relationnel des compétences interculturelles de Imahori et Lanigan, 1989). Bien que ces modèles soient inclus dans différentes catégories, le point commun identifié est le fait que la CI consiste à mobiliser des ressources personnelles pour établir une interaction adéquate avec des personnes appartenant à différentes cultures.

Dans le cadre de notre travail, nous allons considérer trois dimensions principales dans la conceptualisation de la CI - cognitive, affective et comportementale (voir Kohli, 2006; Kim, 2001, et Barmeyer, 2004) (cf. Tableau 1). Par ailleurs, pour opérationnaliser ces dimensions, nous procédons à une triangulation théorique et conceptuelle en tenant compte des travaux de Gudykunst (1993, 2004), Barmeyer (2007), Lázár et collaborateurs (2007), Lussier (2007) et Meunier (2007). Dans la section 4 de cet article, nous expliquerons l'importance de la triangulation pour construire l'outil d'évaluation de la CI.

Tableau 1 : définition opérationnelle des dimensions de la compétence interculturelle

Dimension	Description
Affective (Ressentir)	Avoir une attitude positive ou négative par rapport à quelque chose. Développement d'une conscience (<i>awareness</i>) au sujet des comportements, valeurs et attitudes. <i>Exemples: intérêt, auto-conscience, ouverture d'esprit.</i>
Cognitive (Connaître)	Acte ou état de connaître. Perception claire des faits. Habileté de comprendre de façon active les caractéristiques et les perspectives de personnes appartenant à des cultures différentes. <i>Exemples: attentes, conscience de l'existence de plus d'une perspective.</i>
Comportementale (Agir)	Démontrer la capacité de communiquer et d'interagir de façon appropriée. <i>Exemples: être empathique, tolérer l'ambiguïté, adapter le comportement.</i>

3 La validité et l'évaluation de la compétence interculturelle

Dans la littérature, on peut répertorier plus d'une cinquantaine de questionnaires, échelles ou inventaires d'évaluation de la CI qui peuvent être utilisés pour des fins de sélection, de formation, ou de développement (voir Fantini, 2009; Paige, 2004). Or, l'analyse détaillée des outils qui s'adressent à l'évaluation de la CI en contexte éducatif suscite encore des questionnements.

Bien que des auteurs tels que van de Vijver et Leung (2009), Heine (2008) ou encore Mendenhall et ses collaborateurs (2004) s'intéressent aux procédures méthodologiques d'évaluation de la CI, d'autres aspects méritent d'être questionnés, notamment les caractéristiques qui distinguent l'évaluation de la CI des autres compétences, ainsi que les preuves de validité et de fidélité qui devraient être obtenues dans le processus d'évaluation.

3.1 Analyse des instruments d'évaluation de la CI

En nous basant sur les travaux de Messick (1989) ainsi que de Downing et Haladyna (2009) sur la validité, nous avons analysé 37 instruments d'auto-évaluation. Pour ce faire nous avons construit une grille critériée. La validité de contenu, la qualité psychométrique des instruments, la validité de conséquence, la sensibilité des données ainsi que le contrôle de la désidérabilité sociale sont des aspects constitutifs de cette grille.

Le Tableau 2 qui suit présente une synthèse des éléments principaux de l'analyse, tels que: (1) preuves de validité au niveau théorique et conceptuel, (2) preuves de validité au niveau de la structure interne de l'instrument et (3) preuves de validité au niveau contextuel.

Tableau 2 : Analyse de la qualité des instruments d'auto-évaluation (n=37)

Domaine général de la qualité des instruments	Domaine spécifique de la qualité des instruments	Nombre d'instruments qui rencontrent les critères
1. Preuves de validité aux niveaux théorique et conceptuel	a) Modèle théorique	13
	b) Consultation d'experts	9
	c) Analyses factorielles, corrélations, comparaison avec d'autres instruments	26
2. Preuves de validité au niveau de la structure interne de l'instrument	a) Fidélité; détermination d'autres propriétés psychométriques	29
3. Preuves de validité au niveau contextuel	a) Indications ou comparaison dans plusieurs contextes	11
	b) Réflexions au sujet de l'utilité et pertinence de l'instrument	1

D'un point de vue théorique ou conceptuel (voir Downing et Haladyna, 2009; Sireci, 2009), l'évaluation doit tenir compte d'un référentiel ou d'un modèle théorique consistant. Plus particulièrement, en ce qui concerne l'évaluation de la CI, le plus grand défi méthodologique est au niveau conceptuel (van de Vijver et Leung, 2009). De fait, il s'agit d'un construit multidimensionnel et complexe, car il implique la mesure de composantes cognitives, comportementales et affectives. En conséquence, la définition et l'opérationnalisation de ce construit représentent des tâches exigeantes et complexes pour les personnes qui ont à construire des instruments visant à mesurer des compétences interculturelles.

En outre, les données issues de l'évaluation de la CI peuvent être considérées sensibles. En effet, il s'agit souvent d'identifier des éléments associés à des stéréotypes et/ou des préjugés (voir Lustig et Spitzberg, 1993). Donc, en plus de procéder à des démarches de validation du contenu, il est fondamental de contrôler les biais qui peuvent affecter la validité associée au contexte (Messick, 1993) ou au processus de réponse (Downing et Haladyna, 2009). À ce propos, il sera important de contrôler les biais associés à la

collecte de données, comme par exemple l'influence de l'évaluateur (AEA, 2011) ou encore de contrôler les réponses avec des données concernant la désidérabilité sociale (Johnson et van de Vijver, 2003).

Du point de vue pratique, le Tableau 2 indique qu'uniquement 13 instruments d'auto-évaluation s'appuient sur un ou plusieurs modèles théoriques (exemples: Hammer, 1999; Neville, Lilly, Duran, Lee et Browne, 2000; Van der Zee et Van der Oudenhoven, 2000), et 9 auteurs réfèrent à la consultation d'experts pour valider le contenu (exemples: Bhawuk et Brislin, 1992; Gamst *et al.*, 2004). Bien que ces pratiques idéales ne soient pas adoptées par la majorité des auteurs, un total de 26 instruments d'auto-évaluation ont été validés en procédant à des analyses factorielles, à des analyses corrélationnelles ou encore à travers la comparaison avec d'autres instruments (exemples: Dunbar, 1997; Hammer 2003; Phinney, 1992). On constate que la plupart des auteurs se préoccupent de la validité de contenu. Par contre, nous observons qu'il existe des lacunes au niveau du contrôle des menaces à la validité au niveau individuel et contextuel. En effet, peu d'auteurs mentionnent avoir contrôlé, par exemple, les biais associés à la désidérabilité sociale (exemples: Neville *et al.*, 2000; Ponterotto *et al.*, 2002; Pruegger et Rogers, 1993).

En ce qui concerne les aspects de la validité qui se réfèrent à la structure interne de l'instrument (voir Downing et Haladyna; Sireci, 2009), la plupart des instruments d'auto-évaluation analysés font preuve de bonnes propriétés psychométriques, c'est-à-dire qu'ils ont une valeur élevée de consistance interne (l'ensemble de l'instrument ainsi que de ses dimensions).

Finalement, il est aussi important de tenir compte des conséquences ou implications de l'utilisation d'un instrument (Downing et Haladyna, 2009; Messick, 1993; Sireci, 2009). Cet aspect renvoie vers l'utilité sociale et organisationnelle. En ce qui concerne les objectifs de l'évaluation de la CI, généralement peu d'informations sont fournies en relation à la validité de conséquence (voir D'Andrea, Daniels, et Heck, 1991).

3.2 Réflexion sur le processus d'évaluation de la CI

En plus des aspects spécifiques liés aux instruments pour évaluer la CI, certains auteurs ont indiqué des recommandations plus générales pour évaluer la CI. De ce fait, il est suggéré de trianguler des instruments et de privilégier les méthodologies mixtes (Deardorff, 2006; Fantini, 2009; van de Vijver et Leung, 2009). Effectivement, pour évaluer des construits multidimensionnels et complexes, la méthodologie doit permettre d'obtenir des informations à partir de la perspective des répondants mais aussi à partir d'une perspective externe à ceux-ci, comme à travers le biais d'un observateur. Cette approche permet de corroborer les informations et, par conséquent, de favoriser la validité ainsi que la fidélité des interprétations (voir Kim, 2001). Ces stratégies sont très utiles pour contrôler les biais contextuels ou encore les biais individuels, tels que la désidérabilité sociale. Par ailleurs, puisque les compétences ont un caractère développemental et impliquent la mobilisation de ressources personnelles dans un contexte spécifique (Tardif, 2006), il est recommandé de faire une cueillette systématique des données et procéder à une évaluation par simulation, car cette stratégie permet d'obtenir des données à partir d'une situation réelle (Le Boterf, 2010).

À partir de ces dernières recommandations, une analyse préliminaire de 32 articles scientifiques a été effectuée avec le but d'explorer comment, à l'heure actuelle, les chercheurs et les professionnels mettent en pratique les dispositifs d'évaluation pour mesurer les CI (Simões Forte et Dionne, 2012). Il a été possible de conclure qu'une variété d'instruments est employée dans le cadre de programmes de développement des CI. Toutefois, dans ce contexte particulier, en tenant compte des pratiques favorables à l'évaluation de ces compétences, le choix des méthodes ou des instruments ne s'avère pas tout le temps approprié pour contrer les biais. Par conséquent, la validité et la fidélité des recherches n'est pas toujours bien assurée. À titre d'exemple, peu d'études fournissent des données au sujet du processus d'adaptation

ou de validation des instruments et, finalement, peu d'entre elles tiennent compte du contexte authentique et de la nature sensible des données.

4 Un outil d'évaluation du jugement dans une situation culturellement complexe

Face aux lacunes rencontrées, il semble intéressant d'explorer la contribution d'un instrument qui puisse rendre compte du jugement des répondants dans une situation ambiguë et authentique. Bien que les simulations représentent une option intéressante pour évaluer les compétences (Le Boterf, 2010), elles ne sont pas en mesure de fournir des données standardisées.

À ce propos, les Tests de Concordance de Scripts (TCS) ont été appliqués dans le domaine de la pédagogie médicale (Charlin et Van der Vleuten, 2004; Charlin, Kazi-Tani, Gagnon et Thivierge, 2005) et ensuite adaptés en pédagogie universitaire, plus particulièrement dans la formation à l'enseignement (Dionne et Simões Forte, 2012a). Les principes inhérents à cet instrument partent de la psychologie cognitive et consistent à évaluer le jugement dans une situation complexe, incertaine et authentique (Charlin, Kazi-Tani, Gagnon et Thivierge, 2005). Les résultats des participants sont, ensuite, comparés avec les réponses d'un groupe d'experts.

Plusieurs tentatives d'application de cet instrument en contexte professionnel se sont avérées efficaces. De plus, dans certains cas, les items avaient pour objectif mesurer des composantes liées à la culture (Ross, Uijtdehaage et Lypson, 2010; Uijtdehaage, Henderson et Wilkerson, 2008).

Dès lors, nous proposons d'adapter un TCS au domaine de l'éducation dans le but d'obtenir des informations au sujet de la CI des étudiants en relation aux interactions culturelles vécues en contexte de salle de classe universitaire. Pour ce faire, nous nous sommes appuyés dans la définition opérationnelle des dimensions de la compétence interculturelle définies précédemment.

Nous avons procédé à une triangulation théorique et conceptuelle pour établir un ajustement entre la théorie des scripts, la construction du TCS et le concept de culture. En partant des principes de la théorie de Gudykunst (1993, 2004) où le succès d'une interaction culturelle dépend de la résolution de l'anxiété et la gestion d'une situation incertaine et ambiguë, nous avons complété les propos de cet auteur avec les travaux d'autres auteurs qui avaient déjà défini des manifestations observables des composantes de la CI (voir Barmeyer, 2007; Lázár et collaborateurs, 2007; Lussier, 2007; Meunier, 2007). De cette façon, il a été possible de procéder à une opérationnalisation du concept.

Nous avons retenu trois dimensions principales. Pour la dimension *ressentir*, nous tenons compte de *l'intérêt, de l'auto-conscience et de l'ouverture à de nouvelles informations*. Ces sous-dimensions se manifestent par la volonté d'interagir avec les personnes d'autres cultures, de remettre en question ses propres modes de pensée et de s'intéresser aux valeurs/coutumes des différents groupes ethniques. Ensuite, la dimension *connaître* suppose démontrer la capacité d'anticiper la façon dont les personnes d'autres cultures vont interagir avec nous (sous-dimension *attentes*) et reconnaître que différentes personnes interprètent les comportements de façon différente (sous-dimension *conscience de l'existence de plus d'une perspective*). Finalement, dans la dimension *agir*, nous tenons compte de la *capacité d'être empathique*, autrement dit, la capacité de se mettre à la place des autres et d'agir en tant que médiateur culturel dans des situations difficiles. La *capacité de tolérer l'ambiguïté* est également fondamentale, et consiste à être capable d'analyser et interpréter des événements et situations culturelles complexes, mais également d'être en mesure de réagir confortablement. En définitive, la *capacité d'adapter la communication* consiste à adapter et accommoder le comportement et les objectifs de l'interaction.

La Figure 1 représente un exemple de vignette qui vise à évaluer la sous-dimension *conscience de plus d'une perspective* de la dimension *connaître*. Chaque vignette est associée à des items qui représentent des hypothèses et des conséquences possibles. Par la suite, les répondants doivent indiquer leur choix de réponse à partir de l'échelle de likert présentée.

Vignette (cas, situation ambiguë ou incomplète)

Situation						
Lors d'une rencontre avec vos collègues pour préparer les détails d'une présentation orale pour un cours, vous constatez que l'un de vos collègues, qui est étudiant étranger, adopte une attitude informelle vis-à-vis de la situation ainsi qu'un langage vulgaire non recommandé en contexte professionnel.						
Conséquences						
Vous jugez de la façon suivante :	Suite à votre décision, vous remarquez que :	Vous considérez votre intervention :				
		Inutile	Peu utile	Utile	Très utile	N/A
1. Vous considérez que l'attitude du collègue est inappropriée et vous le réprimandez devant vos collègues.	L'étudiant étranger adopte l'attitude formelle que vous considérez adéquate.					
2. Vous considérez l'attitude inadéquate et vous lui expliquez discrètement comment les étudiants se préparent traditionnellement pour les exposés en groupe dans votre pays d'origine.	Votre collègue adopte une attitude plus formelle vis-à-vis de l'exposé.					
3. Bien que vous considériez l'attitude inadéquate, vous ne faites aucun commentaire.	Votre collègue adopte une posture informelle au cours de l'exposé. La rétroaction du chargé de cours et de vos camarades est positive.					

↑

**Hypothèse
(Plusieurs items
possibles)**

↑

**Effet ou conséquence
possible**

↑

**Réponses possibles
(échelle de likert)**

Figure 1: Vignette – Dimension *connaître* / sous-dimension *conscience de plus d'une perspective*

Chaque vignette présente une situation ambiguë, incertaine ou incomplète. De cette façon, il est possible d'établir des discussions et des débats avec les répondants au sujet de leur jugement et leur choix de réponse. Ainsi, pour chaque item, nous sommes partis de la définition opérationnelle et des comportements ou attitudes considérés adéquats pour créer une grille d'analyse de l'item (cf. Tableau 3). Cette grille peut également servir comme guide de discussion.

Tableau 3 : Éléments de discussion / éléments manquants ou ambigus

Item	Sujets pour discussion/questionnement
1	Est-ce qu'il s'agit d'un exposé final d'un travail d'équipe? Est-ce que le travail est formatif ou sommatif (note finale)? Combien de personnes font partie du groupe? Quelle est l'origine de l'étudiant étranger? Est-ce que la personne est au courant de la culture de travail dans un autre pays? Quelle est la nature de l'intolérance de l'attitude (est-ce qu'il y a des préjugés en ce qui concerne la culture d'origine du collègue qui sont en cause)? Il faut se questionner si dans les faits l'attitude du collègue étranger est inadéquate. Est-ce que les autres collègues jugent de la même façon? (ici la mémoire, les habitudes de travail, les scripts vont influencer le jugement vis-à-vis d'une situation).
2	Ici l'attitude est adéquate dans la mesure où la personne respecte la culture du collègue et assume qu'il n'est pas en mesure d'interpréter les situations et règlements de travail de la même façon que nous. Cela implique la connaissance et la compréhension des procédures traditionnelles de préparation, mais aussi de comprendre la posture de l'autre.
3	Il s'agit d'une attitude passive. Il n'y a pas la motivation de comprendre la perspective de l'autre et on se maintient dans sa propre perspective. Puisqu'il n'y a aucun engagement vers le collègue, nous n'allons pas apprendre, ni lui.

5 Discussion et conclusion

L'administration des TCS peut être suivie de discussions avec les étudiants (Dionne et Simões Forte, 2012b; Ross, Uijtdehaage et Lypson, 2010). Ainsi, nous pensons que cet outil possède de nombreux avantages du point de vue de l'auto-réflexion et du développement de compétences. Par ailleurs, du point de vue pédagogique, il s'agit d'un instrument efficace dans les pratiques d'évaluation formative.

En effet, les discussions permettent d'obtenir des informations concernant le jugement des étudiants par rapport à des situations réelles ou authentiques. Il s'agit d'une activité formative appropriée dans la mesure où elle favorise le questionnement et il existe une implication active des étudiants, ce qui leur permet de faire état de leur progrès (voir Dionne, 2012).

Par ailleurs, l'interaction établie avec les étudiants au cours des débats permet de faire émerger des attitudes ou des conduites considérées adéquates en fonction des exigences de la situation (dans la vignette) (voir Allal, 1999) et, en conséquence, il est possible de favoriser le développement de la CI.

En définitive, nous pensons que l'utilisation d'un TCS pour évaluer la CI présente de nombreux avantages, car elle permet de surmonter les limites des pratiques fréquemment adoptées. En plus de stimuler l'esprit critique et de favoriser la progression des étudiants, il peut être administré dans une variété de contextes et permet d'obtenir des données standardisées. Donc, nous espérons que ces considérations seront utiles pour les professeurs, enseignants et praticiens dans le domaine interculturel.

6 Références

- American Evaluation Association (AEA). (2011). *Public Statement on Cultural Competence in Evaluation*. Fairhaven, MA: Auteur . [en ligne] www.eval.org consulté le 20 juin 2011.
- Allal, L.(1999). Acquisition et évaluation des compétences en situation scolaire. Dans J. Dolz et E. Ollagnier (Dir.), *L'énigme de la compétence en éducation*(pp. 77-94). Bruxelles: De Boeck.
- Barmeyer, C. (2004). Peut-on mesurer les compétences interculturelles ? Une étude comparée France - Allemagne - Québec des styles d'apprentissage. Dans G. Simard et G. Lévesque (Dir.). *La GRH mesurée!* Actes XVe congrès AGRH (pp. 1435-1462), Montréal 2004. [En ligne] http://www.philippepierre.com/lib/exe/fetch.php?media=c_barmeyer_compences_interculturelles.pdf consulté le 25 février 2011.
- Barmeyer, C. (2007). *Management Interculturel et Styles d'Apprentissage : Étudiants et dirigeants en France, en Allemagne et au Québec*. Québec : Presses de l'Université Laval.
- Bennett, M. J. (1986). A developmental approach to training for intercultural sensitivity. *International Journal of Intercultural Relations*, 10, 179-196.
- Bhawuk, D. et Brislin, R. (1992). The measurement of intercultural sensitivity using the concepts of individualism and collectivism. *International Journal of Intercultural relations*, 16, 413-436.
- Byram, M. (1997). *Teaching and assessing intercultural competence*. Clevelton, UK: Multilingual Matters.
- Charlin, B. et Van der Vleuten,C. (2004). Standardized Assessment of Reasoning in Context of Uncertainty. *Evaluation & the Health Professions*, 27(3), 304-319.
- Charlin, B., Kazi-Tani, D., Gagnon, R. et Thivierge, R. (2005). Le test de concordance comme outil d'évaluation en ligne du raisonnement des professionnels en situation d'incertitude. *Revue internationale des technologies en pédagogie universitaire*, 2(1), 22-27.
- D'Andrea, M., Daniels, J. et Heck, R. (1991). Evaluating the impact of multicultural counseling training, *Journal of counseling and development*, 70, 143-150.
- Deardorff, D. (2006). Identification and Assessment of Intercultural Competence as a Student Outcome of Internationalization. *Journal of Studies in International Education*, 10(3), 241-266.
- De Vita, G. (2007). Taking stock: An appraisal of the literature on internationalising HE learning. Dans E. Jones et S. Brown (Dir.). *Internationalising Higher Education* (pp. 154-167). New York: Routledge.
- Dionne, E. (2012). Les systèmes à réponses personnalisées (SRP) : un atout pour faire de l'évaluation formative en salle de classe ? *Mesure et évaluation en éducation*. 35(1), 47-65.
- Dionne, E. et Simões Forte, L. A. (2012a, 4 mai). *Examen au rayon X d'un outil visant à mesurer une compétence avec des télévotants*. Communication présentée dans le cadre du Colloque scientifique international sur les TIC en éducation: bilan, enjeux actuels et perspectives futures, Montréal, Québec.
- Dionne, E. et Simoes Forte, L. A. (2012b). *Le Test de Jugement Professionnel en Évaluation des Apprentissages (TJPEA) : un outil visant à mesurer des compétences avec des télévotants*. Actes du colloque organisé par la Maison des technologies de formation et d'apprentissage Roland-Giguère (MATI Montréal), Université de Montréal. 5 p.
- Downing, S. M., et Haladyna, T. M. (2006). Twelve steps for effective test development. Dans S. M. Downing, et T. M. Haladyna (Éds.), *Handbook of test development* (pp. 3-25). New-York: Routledge.
- Fantini, A. E. (2009). Assessing Intercultural Competence: Issues and Tools. Dans Deardorff, D.K. (Dir.). *The Sage Handbook of Intercultural Competence* (pp.456-476). Thousand Oaks: SAGE.
- Dunbar, E. (1997). The personal dimensions of Difference Scale: Measuring Multi-group identity with four ethnic groups. *International Journal of Intercultural Relations*, 21(1), 1-28.

Actes du 25^{ème} colloque de l'ADMEE-Europe Fribourg 2013 :
Evaluation et autoévaluation, quels espaces de formation

- Fowler, S. et Mumford, S. (1999a). *Intercultural sourcebook (vol.1): cross-cultural training methods*. Yarmouth: Intercultural Press.
- Fowler, S. et Mumford, S. (1999b). *Intercultural sourcebook (vol.2): cross-cultural training methods*. Yarmouth: Intercultural Press.
- Franklin-Craft, A. (2010). *An assessment of the intercultural competence of student administrators*. Thèse de doctorat présentée à Michigan State University. Non publiée.
- Gamst, G., Dana, R. H., Der-Karabetian, A., Aragon, M., Arellano, L., Morrow, G. et Martenson, L. (2004). Cultural competency revised: the California Brief Multicultural Competence Scale. *Measurement and Evaluation in Counseling and Development*, 37, 163-183.
- Gudykunst, W. B. (1993). Toward a Theory of Effective Interpersonal and Intergroup Communication. Dans R. L. Wiseman et J. Koester (Dir.). *Intercultural Communication Competence* (pp. 33-71). London: Sage.
- Gudykunst, W. B. (2004). *Bridging Differences: Effective Intergroup Communication* (4e éd.). Thousand Oaks: SAGE.
- Hammer, M. R. (1999). A Measure of intercultural sensitivity : the intercultural development inventory. Dans S. Fowler et S. Mumford (1999b). *Intercultural sourcebook (vol.2): cross-cultural training methods*. Yarmouth: Intercultural Press.
- Hammer, M. R. (2003). *The Intercultural Conflict Style Inventory [Instrument]*. North Potomac, MD: Hammer Consulting Group.
- Heine, S. J. (2008). *Cultural Psychology*. New York: W.W. Norton.
- Imahori, T. T. et Lanigan, M. L. (1989). Relational model of intercultural communication competence. *Intercultural communication competence*, 13, 269-286.
- Johnson, T. et van de Vijver, F. J. R. (2003). Social desirability in cross-cultural research. Dans J. A. Harkness, F. J. R. van de Vijver et P. P. Mohler (Dir.), *Cross-Cultural Survey Methods* (pp. 195-204). Hoboken, NJ: Wiley-Interscience.
- Kim, Y. Y. (2001). *Becoming Intercultural : an integrative theory of communication and cross-cultural adaptation*. Thousand Oaks: Sage.
- Kohli, H. (2003). *Building an assessment model to measure cultural competence in graduating social work students*. Thèse de doctorat présentée à la graduate school de l'Université de Louisville. Louisville, Kentucky. Non publiée.
- Kohls, L. R. et Knight, J. M. (1994). *Developing intercultural awareness: a cross-cultural training handbook* (2nd ed.). Yarmouth: Intercultural Press.
- Lázár, I., Huber-Krieger, M., Lussier, D., Matei, G. et Peck, C. (2007). *Developing and assessing intercultural communicative competence : A guide for language teachers and teacher educators*. Strasbourg: Conseil de l'Europe.
- Le Boterf, G. (2010). *Repenser la compétence*. Paris : Groupe Eyrolles.
- Lussier, D. (2007, Avril). *Les compétences interculturelles: un référentiel en enseignement et en évaluation*. Présentation au Colloque ALTE/CIEP (Paris, France).
- Lustig, M. W. et Spitzberg, B. H. (1993). Methodological issues in the study of intercultural communication competence. Dans R. L. Wiseman et J. Koester (Dir.). *Intercultural communication competence* (pp. 153-167). Newbury Park: Sage.
- Meade, A. (2010). *Intercultural competence : an appreciative inquiry*. Thèse de doctorat présentée Mendenhall et al
- Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment: *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1993). Validity. Dans R. Linn (Dir.), *Educational Measurement* (pp. 13-103). Washington, DC: American Council on Education.

Actes du 25^{ème} colloque de l'ADMEE-Europe Fribourg 2013 :
Evaluation et autoévaluation, quels espaces de formation

- Mendenhall, M., Stahl, G., Ehnert, I., Oddou, G., Osland, J. et Kuhlmann, T. (2004). Evaluation studies of cross-cultural training programs; a review of the literature from 1988 to 2000. Dans D. Landis, J. Bennett et M. Bennett (Dir.). *Handbook of Intercultural Training* (3^e ed.) (pp. 129-143). Thousand Oaks: SAGE.
- Meunier, O. (2007). *Approches interculturelles en Éducation : Étude comparative internationale*. Lyon : Institut National de Recherche Pédagogique.
- Neville, H. A. , Lilly, R. L., Duran, G., Lee, R. M. et Browne, L. (2000). Construction and initial validation of the Color-Blind Racial Attitudes Scale (CoBRAS), *Journal of Counseling Psychology*, 47, 59-70.
- Paige, R. M. (2004). Instrumentation in intercultural training. Dans D. Landis, J. Bennett et M. Bennett (Dir.). *Handbook of Intercultural Training* (3^e ed.) (pp. 85-122). Thousand Oaks: SAGE.
- Phinney, J. S. (1992). The multigroup Ethnic Identity Measure: A New Scale for use with diverse groups. *Journal of Adolescent Research*, 7, 156-176.
- Ponterotto, J. G., Sanchez, C. M., et Magids, D. M. (1990). *The Multicultural Counseling Awareness Scale (MCAS) : form B, revised self-assessment*. New York: Fordham University.
- Pruegger, V. J. et Rogers, T.B. (1993). Development of a Scale to Measure Cross-Cultural Sensitivity in the Canadian Context. *Canadian Journal of Behavioural Science*, 25(4), 615-621).
- Simões Forte, L.A., et Dionne, E. (2012, 7 au 11 mai). Menaces à la validité et à la fidélité lors de l'évaluation des compétences interculturelles. Communication présentée dans le cadre du 80ième congrès de l'Association francophone pour le savoir - Acfas, Montréal, Québec.
- Sireci, S. G. (2009). Packing and Unpacking Sources of Validity Evidence: History Repeats Itself Again. Dans R. W. Lissitz (Dir.), *The Concept of Validity: Revisions, New Directions and Applications* (pp. 19-37). Charlotte, NC: Information Age Publishing.
- Spitzberg, B. et Changnon, G. (2009). Conceptualizing Intercultural Competence. Dans Deardorff, D.K. (Dir.). *The Sage Handbook of Intercultural Competence* (pp.2-45). Thousand Oaks: SAGE.
- Ross, P. T., Uijtdehaage, S. et Lypson, M. L. (2010). Reflexions on culture: views on script concordance testing. *Medical Education*, 44, 505-506.
- Tardif, J. (2006). *L'évaluation des compétences: documenter le parcours de développement*. Montreal : Chenelière Education.
- Tomalin, B. (2009). Applying the principles: Instruments for intercultural business training. Dans A. Feng, M. Byram, et M. Fleming (Dir.), *Becoming interculturally competent through education and training* (pp. 115-131). Toronto: Multilingual Matters.
- Uijtdehaage, S. Henderson, P. et Wilkerson, L. (2008, mars). *An instrument for assessing knowledge of cross-cultural care*. Communication présentée à la 13e Conférence Internationale de Compétence Clinique (the 13th Ottawa International Conference on Clinical Competence), Melbourne, Australie.
- Van Oudenhoven, J. et Van der Zee, K. (2002). Predicting multicultural effectiveness of international students: the Multicultural Personality Questionnaire. *International Journal of Intercultural Relations*, 26, 679-694.
- van de Vijver, F. J. R. et Leung, K. (2009). Methodological Issues in researching Intercultural Competence. Dans Deardorff, D.K. (Dir.). *The Sage Handbook of Intercultural Competence* (pp.404-418). Thousand Oaks: SAGE.