

## L'entrevue cognitive : les apports et les limites dans la validation d'un questionnaire

## The cognitive interview: contributions and limitations in questionnaire validation

## A entrevista cognitiva: contribuições e limites na validação de um questionário

**Eliane Dulude**

ID ORCID : 0000-0001-8930-4956

*Université d'Ottawa*

**Carole Fleuret**

ID ORCID : 0000-0002-1586-5109

*Université d'Ottawa*

**Ludivine Bodar**

*Université d'Ottawa*

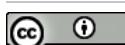
**Marie-Pier Bastien**

*Université d'Ottawa*

**Marie-Hélène Ayotte**

ID ORCID: 0000-0001-7863-2567

*Université du Québec en Abitibi-Témiscamingue*



MOTS CLÉS: adaptation transculturelle, enseignement secondaire, entrevue cognitive, rôle des intervieweurs, validation

*Les politiques d'imputabilité liées aux résultats des tests influencent les pratiques des intervenants scolaires du secondaire, soit les enseignants et les conseillers pédagogiques. Un questionnaire a été adapté afin de mesurer les facteurs sociocognitifs influençant leurs changements pédagogiques en Ontario et au Québec. Afin de l'adapter de façon transculturelle et de le valider, des entrevues cognitives ont été menées auprès de neuf participants par quatre intervieweurs. Une grille d'observation a été créée selon deux types de protocoles utilisés: le protocole à voix haute et les relances. Cette étude se base, d'une part, sur des données quantitatives telles que la durée des entrevues, le nombre et le type de relances et, d'autre part, sur des données qualitatives telles que la qualité et la variété des relances. Nos résultats soulignent ainsi qu'il existe une variation à l'intérieur de ces protocoles qui peuvent s'inscrire sur un continuum en catégorisant les intervieweurs selon trois rôles: passif, modéré et actif.*

KEY WORDS: adaptation, cognitive interview, high school teachers, interviewers' role, validation

*Accountability policies linked to test scores influence secondary intervenors' practices (e.g. teachers and educational consultants). A questionnaire was adapted to measure socio-cognitive factors influencing pedagogical change among secondary school intervenors in Ontario and Quebec. To adapt the questionnaire transculturally and validate it, cognitive interviews were conducted with nine participants by four interviewers. An observation grid was created based on two types of protocols: the think-aloud protocol and the prompting protocol. This study first relies on quantitative data such as interview duration, the number and type of probes. The study, then, uses qualitative data to investigate the quality and variety of their probes. Our results show that there is variation within these protocols, which can be categorized on a continuum into three interviewer roles: passive, moderate and active.*

**PALAVRAS-CHAVE:** adaptação transcultural, ensino secundário, entrevista cognitiva, papel dos entrevistadores, validação

*As políticas de responsabilização ligadas aos resultados dos testes influenciam as práticas dos intervenientes escolares do ensino secundário, nomeadamente professores e consultores pedagógicos. Um questionário foi adaptado para medir os fatores sociocognitivos que influenciam as mudanças pedagógicas destes profissionais em Ontário e Québec. Para garantir uma adaptação transcultural e validar o instrumento, foram realizadas entrevistas cognitivas com nove participantes por quatro entrevistadores. Foi criada uma grelha de observação com base em dois tipos de protocolos utilizados: o protocolo de pensar em voz alta e os estímulos verbais. Este estudo baseia-se, por um lado, em dados quantitativos, como a duração das entrevistas e o número e o tipo de estímulos verbais, e, por outro lado, em dados qualitativos, como a qualidade e a variedade desses estímulos. Os resultados destacam variações dentro desses protocolos, que podem ser categorizadas num continuum com três papéis distintos dos entrevistadores: passivo, moderado e ativo.*

## Problématique

L'entrevue cognitive est une méthode utilisée pour recueillir des informations verbales lors de tests préliminaires et lors de la validation d'un nouveau questionnaire ou d'un questionnaire adapté (Drennan, 2003). Cette approche permet d'explorer la compréhension des participants sur les items, de détecter les problèmes de traduction ou de langage et d'améliorer la validité et la fiabilité des items en fonction du contexte et de la population ciblée (Ridolfo & Schoua-Glusberg, 2011; Sopromadze & Moorosi, 2017). Différentes techniques, telles que la reformulation, le reflet et le silence, sont employées pour obtenir divers types de données auprès des participants (p. ex., compréhension de concepts, vérification des informations) (Priede et al., 2014) et pour permettre à l'intervieweur d'adopter différents rôles afin d'obtenir ces informations (Beatty & Willis, 2007).

Certaines études suggèrent que, pour adapter et pour valider efficacement un questionnaire, les intervieweurs doivent être en mesure de produire des informations verbales pertinentes. Cela permet d'évaluer les problèmes de compréhension et de détecter les biais ou les erreurs potentielles de mesure (Priede & Farrall, 2011; Sopromadze & Moorosi,

2017). L'étude de Priede et al. (2014), qui a comparé les techniques pour susciter la réflexion à voix haute (*think aloud*) à celles des relances (*probing*), montre que ces dernières identifient davantage de problèmes dans les items du questionnaire, tandis que la réflexion à haute voix est plutôt adaptée pour examiner la compréhension des participants des concepts sous-jacents. De plus, les participants aux protocoles de réflexion à haute voix sont plus susceptibles de ne pas répondre à au moins un item, en particulier pour les questions complexes et longues, comparativement à ceux qui participent aux protocoles de relance (Priede et al., 2014). L'étude de Priede et al. (2014) a analysé près de 3 000 relances pour déterminer dans quelle mesure il est pertinent de les varier afin de permettre aux participants de comprendre les items d'un questionnaire. Les résultats révèlent que les relances spontanées sont les moins productives en ce qui concerne les données collectées. Il semble préférable de combiner des questions anticipées, émergentes et conditionnelles lors des entretiens cognitifs pour obtenir des informations plus pertinentes. Cependant, il n'y a pas de consensus sur les meilleures techniques à utiliser, car leur efficacité dépend des populations et des contextes étudiés.

D'autres études se concentrent spécifiquement sur les comparaisons transculturelles. Elles étudient comment les contextes socioculturels influencent le processus de réponse des participants aux enquêtes. Elles examinent également les effets des interactions entre les intervieweurs et les interviewés (Ridolfo & Schoua-Glusberg, 2011; Willis & Miller, 2011; Willis, 2015). Dans son étude auprès d'adultes hispanophones, García (2011) a mis en lumière deux principaux défis : 1) des problèmes de compréhension linguistique (en anglais et en espagnol) et 2) des difficultés liées au niveau de langage utilisé (formel ou familier). Willis (2015) a effectué une synthèse des connaissances sur l'entretien cognitif transculturel à partir de 32 études recensées, en se centrant sur leurs objectifs, leurs procédures et leurs résultats. Les résultats ont mis en évidence l'impact des différences culturelles et linguistiques ainsi que l'importance de la sélection et de la formation des intervieweurs pour mener à bien les entretiens cognitifs. L'étude de Sopromadze et Moorosi (2017) a cherché à adapter un questionnaire bilingue portant sur la perception des pratiques émotionnelles de leadership en milieu universitaire dans les contextes anglais et géorgien en évaluant la validité cognitive, l'équivalence sémantique et l'équivalence conceptuelle de leurs mesures. Les résultats aux entrevues cognitives montrent, entre autres, que la compréhension des items relatifs aux valeurs

sociétales varie entre les échantillons des deux contextes d'application. En effet, les participants anglais s'interrogeaient à savoir si l'expression « la plupart des gens » faisaient référence au groupe ethnique prédominant, alors que les participants géorgiens l'associaient aux Géorgiens.

### ***Les contextes de comparaison culturelle et politique***

Notre étude élargie s'est intéressée à la comparaison des réalités des intervenants scolaires dans les provinces de l'Ontario et du Québec en fonction des enjeux élevés liés aux tests (p. ex., exigences de réussite pour l'obtention du diplôme), par rapport à ce qu'ils représentent pour eux. Elle s'est également intéressée à la comparaison de leurs contextes socio-culturels susceptibles d'expliquer les variations. Les études canadiennes recensées sur la manière dont l'utilisation des résultats aux examens provinciaux et les mesures incitatives (p. ex., exigences de réussite aux tests pour l'obtention du diplôme, mécanismes de suivi des résultats scolaires) révèlent une certaine réactivité chez les intervenants (Copp, 2016, 2018; Maroy et al., 2017; Yerly, 2017). Celles-ci indiquent, entre autres, que les intervenants scolaires du secondaire, dans les provinces où les examens de fin d'études sont certificatifs, sont plus susceptibles d'utiliser des pratiques pédagogiques centrées sur les tests que ceux des provinces qui n'ont pas de tels examens. De plus, elles ont montré que, même si ni les écoles ni les intervenants ne font l'objet de sanctions en cas de résultats faibles, les enjeux associés à la divulgation des résultats au grand public et de l'opinion de celui-ci à la suite de leur publication semblent exercer une pression accrue sur les pratiques enseignantes.

Ces études ont aussi montré qu'il existe une variation provinciale qui semble être une piste explicative importante sur le degré de réactivité et sur les comportements adoptés par les intervenants scolaires (Copp, 2016, 2018; Maroy et al., 2017; Yerly, 2017). En effet, certaines études suggèrent que des mesures incitatives politiques propres à chaque province risquent d'orienter les pratiques vers des activités de préparation aux tests (Copp, 2018). Or, force est de constater que la comparaison entre les provinces exige une attention particulière aux facteurs socioculturels et politiques afin de connaître leurs répercussions sur les représentations et sur les types de changements apportés par les intervenants. Les mesures sélectionnées reposent sur les constats de réformes antérieures au Canada qui estiment que, pour induire des changements dans la pratique à l'échelle systémique, il faut une combinaison de pression induite par des mesures incitatives et

de soutien adapté (Fullan, 2000). Nous avons donc choisi les provinces de l'Ontario et du Québec en fonction des mécanismes incitatifs politiques variés associés aux tests (p. ex., soutien par les ministères, instances intermédiaires et directions, nombre d'évaluations certificatives) et des enjeux susceptibles d'être plus élevés, compte tenu de ce que ces mesures représentent pour les intervenants. La comparaison des éléments culturels et politiques entre les provinces, dont les mesures incitatives politiques et le soutien offert, montre des conceptions divergentes quant à la manière d'apporter des changements dans la pratique. C'est pourquoi nous avons d'abord examiné le contexte, tel que le soutien offert par les ministères, et d'autres mesures incitatives politiques, telles que les exigences de réussite aux tests pour l'obtention du diplôme et le nombre d'évaluations certificatives, afin de s'interroger sur les différences de facteurs culturels et politiques entre l'Ontario et le Québec, et ce, autant lors des entrevues cognitives que pour l'administration du questionnaire à plus grande échelle.

En Ontario, le ministère de l'Éducation se concentre sur trois priorités : l'amélioration des résultats scolaires en mathématiques et en littératie ainsi que l'augmentation des taux de diplomation. L'Office de la qualité et de la responsabilité de l'éducation (OQRE), un organisme paragouvernemental, est chargé d'administrer et de superviser les tests standardisés obligatoires en mathématiques et en littératie (par exemple, écriture et lecture). L'OQRE est également responsable du suivi des progrès auprès des conseils scolaires et des écoles, ainsi que de l'obligation de soumettre un plan d'amélioration scolaire incluant les résultats scolaires (Anderson & Ben Jaafar, 2003). De plus, le Ministère a offert un soutien gouvernemental intitulé : *Gestion de l'information pour l'amélioration du rendement des élèves (GIARE)* qui comprenait deux composantes importantes : 1) le renforcement des capacités locales pour utiliser les données lors de la prise de décision et 2) la mise en place de systèmes de gestion des données afin de les colliger, de les stocker et de les analyser (Dunn et al., 2013). Quant aux évaluations certificatives, seulement deux sont administrées au secondaire : les tests en 9<sup>e</sup> année (mathématiques), qui sont des épreuves uniques et qui comptent pour jusqu'à 10 % du résultat final des élèves et le test de littératie, en 10<sup>e</sup> année, dont la réussite est nécessaire pour l'obtention du diplôme d'études secondaires (Winton & Milani, 2017).

Au Québec, le Ministère se concentre sur cinq buts ministériels : 1) l'amélioration de la diplomation et de la qualification avant 20 ans, 2) l'amélioration de la langue française, 3) l'amélioration de la persévérance

et de la réussite scolaire chez certains groupes ciblés, particulièrement les élèves handicapés ou en difficulté d'adaptation ou d'apprentissage, 4) le développement d'un environnement sain et sécuritaire dans les établissements et 5) l'augmentation du nombre d'élèves de moins de 20 ans en formation professionnelle. Le Ministère a la responsabilité de deux types d'épreuves : les épreuves obligatoires et les épreuves uniques. Ainsi, les épreuves obligatoires sont des épreuves de régulation qui vérifient l'atteinte des exigences de certains programmes. Dans le cas des évaluations certificatives, il existe cinq épreuves uniques qui comptent pour 50 % de la note finale des élèves de 4<sup>e</sup> et de 5<sup>e</sup> secondaire dans les matières dont la réussite est exigée pour l'obtention du diplôme d'études secondaires.

Bien que les entrevues cognitives permettent de comprendre la façon dont les personnes appréhendent les items d'un questionnaire, le processus d'adaptation et de validation de l'explicitation de leur pensée est complexe. L'utilisation de ces entrevues nécessite, en effet, une réflexion approfondie sur l'influence du type de procédure employée (p. ex., réflexion à voix haute, relance) et des techniques utilisées (p. ex., reformulation, paraphrase) sur la qualité des réponses obtenues. Cela comprend la prise en compte du contexte ainsi que les caractéristiques des intervieweurs et des participants qui façonnent ce processus. À la lumière des études recensées et des facteurs socioculturels et politiques propres au Québec et à l'Ontario, il semble pertinent d'analyser la manière dont les participants dans ces deux contextes comprennent les éléments liés aux pratiques pédagogiques des enseignants lors de la préparation aux examens. L'intérêt de notre étude est de décrire le rôle des intervieweurs dans le processus d'adaptation transculturelle et dans la validation des items du questionnaire. Dans le cadre d'entrevues cognitives, ces items cherchaient à mesurer, chez les intervenants scolaires, les facteurs sociocognitifs qui les prédisposent à changer leurs pratiques en fonction des résultats de leurs élèves aux tests provinciaux.

## Cadre théorique

La validité joue un rôle central dans la conception des items ou des échelles (Castillo-Diaz & Padilla, 2013) ainsi que lors de la traduction et de l'adaptation des mesures (Grondin et al., 2015). Bien que cette notion soit constamment utilisée en recherche, elle semble complexe (Loye, 2018). Loye (2018) explique que ce qui est intéressant dans une épreuve de mathématiques n'est pas le fait de donner une note, mais de savoir si la note

correspond à ce que l'épreuve voulait évaluer. Son utilité, quant à elle, permet de garantir l'utilisation d'un instrument peut être généralisée. Cette généralisation demeure cependant limitée à une population donnée qui est représentée par l'échantillon interrogé lors du processus de recherche (Ajar et al., 1983). Toutefois, la validité est souvent cantonnée aux propriétés psychométriques de l'instrument sans prendre en compte la population et le contexte d'application. Un questionnaire mal adapté au contexte dans lequel il est administré peut occasionner des problèmes de fiabilité et de validité (Gudmundsson, 2009). En effet, selon Gana et al. (2021), l'instrument peut présenter trois principaux biais : 1) le biais de construit (compréhension du sujet selon la culture), 2) le biais de méthodologie (mode d'administration, choix de l'outil, etc.) et 3) le biais d'items (signification d'un item d'une culture à une autre). Une traduction ne semble pas suffisante pour un instrument validé dans une autre langue.

La recherche de validité commence dès la conception ou l'adaptation d'un instrument (Loye, 2018) et suit une démarche qui permet de réduire les biais qui pourrait l'affecter. En ce sens, la validité prend une place importante dans la traduction et dans l'adaptation transculturelle parce que la démarche exige de considérer à la fois les aspects de traduction et ceux qui sont liés à l'adaptation contextuelle (culturelle) (Corbière & Fraccaroli, 2020). Selon Corbière et Fraccaroli (2020), cette démarche s'effectue en six étapes. Premièrement, certaines dimensions préalablement validées par les chercheurs font l'objet d'une traduction initiale (de la langue d'origine à la langue ciblée). Deuxièmement, une version de l'instrument comprenant la version originale et les subséquentes est nécessaire. Troisièmement, un professionnel procède à une traduction inversée, c'est-à-dire de la langue ciblée vers la langue d'origine afin de vérifier la validité de la traduction et la fidélité du contenu des items. Quatrièmement, un comité d'experts (équipe de chercheurs, de traducteurs et un participant de la population cible) est mis sur pied pour effectuer la révision des items. La révision par le comité permet de respecter certains critères de clarté, de précision et de simplicité des consignes et des items ainsi qu'un vocabulaire accessible pour les participants ciblés. Lauzier et al. (2021) évoquent les spécificités linguistiques et culturelles d'un instrument de mesure. Gana et al. (2021) expliquent qu'« adapter un test exige de le rendre équivalent – au niveau psychologique – à la version originale » (p. 255). La traduction littérale ne suffit pas, car il est nécessaire d'adapter l'instrument sur les plans culturels et sémantiques. Cinquièmement, le pilotage de la nouvelle

version de l'instrument auprès d'un petit groupe de participants permet de tester non seulement la traduction, mais également l'adaptation des items au contexte. Cette étape constitue le dernier rempart avant l'administration du questionnaire et l'analyse préliminaire des items (Gana et al., 2021). Le fait d'avoir intégré l'aspect transculturel dans les étapes précédentes permet d'éviter une mauvaise compréhension ou une mauvaise interprétation des items et de choisir la méthode appropriée selon la population définie. Cette cinquième étape est donc cruciale dans la démarche de validation. À cet égard, l'entretien cognitif semble l'un des moyens de tester l'adaptation du questionnaire (Gana et al., 2021). La sixième et dernière étape porte sur la validation de l'instrument de mesure auprès d'un plus grand échantillon (Corbière & Fraccaroli, 2020).

## 2.1 Les entrevues cognitives

Les entrevues cognitives prennent appui sur les travaux menés en psychologie cognitive et des méthodes de sondage afin de comprendre comment s'opère le traitement de l'information, ce qui inclut l'attention, la reconnaissance des mots, la mémoire, le traitement du langage, la résolution de problèmes et le raisonnement (Drennan, 2003). Autrement dit, la verbalisation des items y est considérée comme une représentation du traitement de l'information par les participants. Les entrevues cognitives sont donc une méthode qualitative qui explorent la manière dont les personnes traitent mentalement les informations lorsqu'elles répondent à des questionnaires (García, 2011). Ces entrevues permettent de recueillir des informations verbales supplémentaires sur les réponses (Beatty & Willis, 2007).

Les entrevues cognitives peuvent reposer sur au moins deux protocoles : 1) la réflexion à voix haute (*think aloud*) et 2) la relance (*probing*). Le protocole de réflexion à voix haute vise à faire émerger les réflexions lors d'activités où des participants sont invités à exprimer leurs pensées en lisant les items et en sélectionnant leurs choix, ce qui limite le rôle de l'intervieweur (Beatty & Willis, 2007). Celui-ci devient un facilitateur de la verbalisation et intervient le moins possible dans la production de ces informations. En ce sens, l'interviewé est convié à lire à voix haute et à expliciter le fil de ses pensées. Le protocole de relance, quant à lui, permet une présence accrue de l'intervieweur qui pose des questions directes à l'interviewé (Beatty & Willis, 2007). Ce protocole est favorable pour atténuer la gêne que peuvent ressentir certains participants et permet, le cas échéant, de compenser la faible quantité des données recueillies. Selon ces auteurs, parallèlement au protocole de réflexion à voix haute, et sans interférer, des praticiens ont ajouté cette dimension au processus de collecte de données.

Cet article vise à décrire les apports et les limites de l'utilisation des entrevues cognitives en s'attardant aux rôles des intervieweurs. Nous cherchions plus particulièrement à identifier et à comparer les types de techniques, les indicateurs comportementaux des intervieweurs ainsi que les réactions des interviewés qui facilitent ou freinent le processus d'adaptation transculturelle et de validation des items dans les contextes québécois et ontarien.

## Méthodologie

La démarche de traduction et d'adaptation transculturelle du questionnaire pour faciliter les comparaisons de similitudes et de différences interindividuelles dans les deux provinces et dans des contextes différents s'est déroulée en trois étapes (Bouchard, 2000; Sabourin et al., 2000). Dans la première étape, nous avons effectué une recension des écrits pour identifier les variables sociocognitives qui prédisposent les intervenants scolaires à changer leur pratique à la suite des résultats obtenus aux examens ministériels. Nous avons retenu cinq échelles : 1) les attitudes, 2) l'attribution causale, 3) l'influence sociale, 4) le sentiment d'efficacité personnelle et 5) les types de changements pédagogiques.

Pour la seconde étape, nous avons traduit trois des cinq échelles de l'anglais au français. Nous avons fait appel à un comité d'experts composé de trois personnes, principalement des chercheurs, pour faire une première lecture complète du questionnaire. Le comité s'est rencontré à deux reprises pour réviser les items. La révision par les pairs a permis d'assurer le respect de certains critères, tels que la clarté, la précision et la simplicité des items ainsi que la pertinence du langage en fonction des participants ciblés (p. ex., âge, niveau de scolarité) et de leur contexte (Ontario et Québec) (Gagnon et al., 2018). Nous cherchions ainsi à cibler les items pouvant susciter des biais de compréhension ou présentant un degré élevé de polarisation. Nous avons apporté les premières modifications en ajoutant trois groupes sociaux, soit les parents, les médias et les conseils/commissions scolaires à l'échelle de l'influence sociale de Rosenwajn et Dumay (2016) en fonction des résultats recensés dans les études canadiennes (Copp 2016, 2018) et des groupes susceptibles d'exercer des pressions quant à leur responsabilité à l'égard des résultats scolaires. Nous avons également ajouté des notes explicatives à certains endroits du questionnaire, sous recommandation des experts.

Tableau 1  
*Le choix des échelles, de leur construit théorique et de leur version originale*

<b>Échelles de mesure</b>	<b>Construit théorique</b>	<b>Version originale</b>
<i>Conceptions and Practices of Formative Assessment de Yang et Cheng (2015)</i>	Mesurer l'attitude affective et l'attitude opératoire en tenant compte à la fois de la manière dont les enseignants sont prédisposés émotionnellement à agir en fonction des résultats et à évaluer les conséquences de leurs actions.	<b>Contexte d'application :</b> Attitudes, intentions et pratiques d'évaluation formative des enseignants dans le cadre d'une réforme <b>Population :</b> enseignant du primaire <b>Langue :</b> anglais <b>Pays :</b> Hong Kong
<i>Locus of control de Guskey et Passaro (1994)</i>	Mesurer les croyances selon lesquelles les enseignants sont capables d'influencer la réussite des élèves et d'en contrôler les résultats, ainsi que l'attribution des résultats à des causes externes (p. ex., statut socioéconomique).	<b>Contexte d'application :</b> les perceptions de l'influence personnelle, du pouvoir et de l'impact sur les situations d'enseignement et d'apprentissage <b>Population :</b> enseignants chevronnés et enseignants en formation <b>Langue :</b> anglais <b>Pays :</b> États-Unis
<i>Data-driven decision-making self-efficacy de Dunn, Airola, Lo et Garrison (2013)</i>	Mesurer l'analyse et l'interprétation des données, ainsi que l'application des données dans la pratique. Elle vérifie les croyances qu'ont les enseignants en leurs habiletés à analyser et à interpréter des résultats ainsi qu'à faire des liens avec leur pratique pour améliorer l'apprentissage des élèves.	<b>Contexte d'application :</b> le sentiment d'efficacité et l'anxiété dans la prise de décision éclairée par les résultats scolaires <b>Population :</b> enseignants de la maternelle au secondaire <b>Langue :</b> anglais <b>Pays :</b> États-Unis

Tableau 1

*Le choix des échelles, de leur construit théorique et de leur version originale (suite)*

<b>Échelles de mesure</b>	<b>Construit théorique</b>	<b>Version originale</b>
<i>L'influence sociale de Rosenwajn et Dumay (2016)</i>	Mesurer la perception que les enseignants ont de l'importance accordée par leurs collègues, leur direction et les parents aux résultats des examens ministériels.	<b>Contexte d'application :</b> les processus cognitifs intervenant à l'égard des résultats scolaires dans le cadre d'une nouvelle évaluation externe certificative standardisée obligatoire à la fin du premier cycle du secondaire <b>Population :</b> enseignants préparant les élèves aux examens certificatifs de la fin du premier degré de l'enseignement secondaire <b>Langue :</b> Français <b>Pays :</b> Belgique francophone
<i>Types de changements pédagogiques de Rosenwajn et Dumay (2016)</i>	Mesurer les types de changements pédagogiques rapportés par les enseignants à la suite des résultats aux examens provinciaux.	<b>Contexte d'application :</b> les processus cognitifs intervenant à l'égard des résultats scolaires dans le cadre d'une nouvelle évaluation externe certificative standardisée obligatoire à la fin du premier cycle du secondaire <b>Population :</b> enseignants préparant les élèves aux examens certificatifs de la fin du premier degré de l'enseignement secondaire <b>Langue :</b> Français <b>Pays :</b> Belgique francophone

À la troisième étape, pour répondre à notre objectif initial, soit l'adaptation transculturelle des construits et la validation auprès d'intervenants du secondaire des matières scolaires évaluées par les examens provinciaux (mathématiques et français) dans les deux provinces, nous avons choisi les entrevues cognitives comme moyen d'étudier la manière dont ces intervenants traitent mentalement les informations en répondant au questionnaire. Nous cherchions ainsi à cerner les particularités des contextes de comparaison et d'application du questionnaire ainsi que les distinctions quant au langage, à la culture et au contexte politique dans lequel s'effectue la

passation des examens ministériels (Corbière & Fraccaroli, 2020). Nous avons ainsi cherché à adapter les échelles de mesure provenant de la Belgique francophone, de Hong Kong et des États-Unis en établissant des équivalences auprès des populations ciblées lors des entrevues cognitives.

### ***La collecte de données***

Avant de commencer la collecte de données, nous avons formé les intervieweurs. Deux premiers intervieweurs ont testé des techniques (p. ex., abstention, paraphrase, relance) pour réaliser les entrevues cognitives. Après chaque entrevue, ils se sont donné de la rétroaction (une personne observait à l'aide d'une grille préliminaire permettant de détecter les difficultés, l'autre effectuait l'entrevue). Ces premières entrevues ont duré environ une heure chacune et se sont déroulées dans un lieu physique isolé. À partir de ce premier ensemble d'entrevues, nous avons conçu des outils pour soutenir la démarche d'entrevue, incluant des consignes à suivre avant (p. ex., but de l'entrevue, la démarche d'extériorisation de la pensée) et pendant les entrevues (p. ex., des exemples de techniques à utiliser). Deux nouveaux intervieweurs se sont ajoutés à l'équipe dans le contexte de la pandémie de COVID-19. Leur formation s'est donc basée sur les enregistrements initiaux et ils ont reçu une explication de la démarche et de l'outil servant à la réalisation des entrevues par vidéoconférence. Nous offrons un bref portrait des caractéristiques démographiques et socio-culturelles des intervieweurs d'une étude pour comprendre le contexte et l'impact potentiel de ces caractéristiques lors des entretiens cognitifs.

En ce sens, la durée de la formation, le degré d'expérience des intervieweurs par rapport à l'entrevue cognitive ainsi que les conditions de formation ont varié. Au total, quatre intervieweurs ont effectué les entrevues selon deux modes de passation : en personne et en virtuel. Le passage au mode virtuel a été nécessaire dans le contexte de la pandémie de COVID-19.

### ***L'analyse des données***

Les entrevues cognitives nécessitent de s'engager dans un processus itératif, c'est-à-dire de mener quelques entrevues, puis d'en faire le traitement et de poursuivre afin d'apporter les ajustements nécessaires en cours d'adaptation. Ceci permet d'améliorer à la fois la formulation des items et leur compréhension au sein des populations ciblées (Desimone & Le Floch, 2004). Nous avons donc analysé le contenu dès les premières

Tableau 2  
*Les caractéristiques démographiques et socioculturelles des intervieweurs*

Intervieweurs	Âge	Lieu de résidence	Origine ethnoculturelle	Expériences préalables avec la méthode
Étudiant-Doctorat	33 ans	Région urbaine, Ontario	Africaine	<ul style="list-style-type: none"> <li>• Observations préalables d'entretiens</li> <li>• Rétroactions à partir d'une grille préliminaire</li> </ul>
Étudiant-Maîtrise	24 ans	Région rurale, Québec	Nord-américaine	<ul style="list-style-type: none"> <li>• Écoute d'enregistrements d'entretiens initiaux</li> <li>• Lectures des consignes avant et après les entretiens</li> <li>• Séance d'explications</li> </ul>
Professeur 1	54 ans	Région périurbaine, Québec	Européenne	<ul style="list-style-type: none"> <li>• Séance d'explications</li> <li>• Lectures des consignes avant et après les entretiens</li> </ul>
Professeur 2	41 ans	Région urbaine, Ontario	Nord-américaine	<ul style="list-style-type: none"> <li>• Entretiens préliminaires et expérimentation de techniques variées</li> <li>• Échanges des observations à partir d'une grille préliminaire</li> </ul>

entrevues, afin d'identifier les problèmes et d'ajuster les items. Nous avons choisi l'échelle de Dunn et al. (2013), conçue par des chercheurs offrant des formations sur l'utilisation des résultats scolaires pour améliorer la pratique des intervenants scolaires. Ce choix était guidé par le contexte politique ontarien et par la mise en place d'une stratégie provinciale s'appuyant sur les résultats scolaires des élèves pour former les enseignants et modifier leurs pratiques (Dunn et al., 2013). Cette comparaison des contextes semblait pertinente pour mesurer de quelle manière les intervenants des deux provinces interprètent des résultats scolaires pour améliorer leur pratique. Or, nos premières entrevues cognitives nous ont révélé que certains intervenants ontariens ne reçoivent pas les résultats, à moins que la performance moyenne de leurs groupes soit faible. Cela nous a amenées à ajouter deux questions, à savoir : les intervenants reçoivent-ils les résultats de leurs élèves ? Et si oui, sous quelle forme les reçoivent-ils ? En tout,

nous avons produit sept versions du questionnaire afin de le formater, de déterminer l'ordre des échelles et d'adapter les items. Ces entrevues ont été filmées, enregistrées et transcrrites pour des fins d'analyse.

Nous avons conçu une grille d'observation à partir des techniques recensées dans plusieurs études (Beatty & Willis, 2007; García, 2011). Cette grille nous a permis d'analyser des données verbales et non verbales des intervieweurs et des interviewés afin d'observer les éléments qui facilitent ou qui contraignent l'adaptation et la validation lors des entrevues cognitives. Cette grille comporte deux dimensions : 1) les procédures de protocole à voix haute et 2) les procédures de la relance. Nous y avons intégré des catégories liées aux rôles des intervieweurs et des informations obtenues par les interviewés lors de ces entrevues.

### ***Les participants***

L'échantillon est constitué de participants recrutés sur une base volontaire dans des écoles secondaires du Québec et de l'Ontario. Ils ont été sélectionnés en fonction des critères d'inclusion suivants : 1) enseignants en fonction ou qui ont enseigné pendant au moins deux ans une matière évaluée par les tests provinciaux (p. ex., mathématiques, français), 2) enseignants d'un établissement secondaire francophone ou 3) conseillers pédagogiques qui accompagnent les enseignants dans ces matières. Dans le tableau 3, nous incluons les caractéristiques démographiques et socioculturelles des participants.

En tout, nous avons interviewé neuf participants provenant des deux provinces.

## **Résultats**

Nos résultats ont d'abord fait ressortir deux échelles plus problématiques que les autres : l'attribution causale et l'attitude. Ces échelles ont soulevé des problèmes de compréhension des items et des éléments contextuels qu'il fallait mieux situer dans les contextes des provinces. Or, dans le processus d'adaptation transculturelle et de validation, nous avons identifié certains apports et certaines limites des entrevues cognitives attribuables à au moins trois dimensions à prendre en compte lors de cette démarche, soit le rôle des intervieweurs, la qualité et la variété de leurs techniques ainsi que leur capacité à obtenir des retours pertinents des interviewés en fonction des caractéristiques socioculturelles.

**Tableau 3**  
***Les caractéristiques démographiques et socioculturelles des participants***

Participants	Sexe	Matières enseignées	Nombre d'années d'expérience en enseignement	Niveau de scolarité	Lieu de résidence
PS1	F	Français et histoire au secondaire	11 ans	Baccalauréat en enseignement	Région périurbaine, Québec
PS5	M	Mathématiques	5 ans	Maîtrise en éducation	Région urbaine, Ontario
PS6	F	Français		Maîtrise en éducation	Région urbaine, Ontario
PS7	M	Mathématiques et français	16 ans	Maîtrise	Région urbaine, Ontario
PS8	F	Français	19 ans	Baccalauréat en lettres françaises, baccalauréat en éducation secondaire, profil histoire et français.	Région urbaine, Ontario
PS9	M	Français	11 ans	Maîtrise en sociologie de l'éducation	Région urbaine, Ontario
PS10	F	Français	Moins de 5 ans	Baccalauréat en enseignement	Région urbaine, Ontario
PS11	Ne s'applique pas à moi	Mathématiques	10 ans	Maîtrise	Région urbaine, Ontario
PS12	F	Français	6 ans	Maîtrise en éducation	Région périurbaine, Québec

### ***Le rôle des quatre intervieweurs***

Pour caractériser le rôle des intervieweurs, nous avons analysé le nombre d'interventions par entrevue et identifié les types de techniques utilisées (p. ex., paraphrase, silence, relance, etc.) comme l'illustre le tableau 4 ci-dessous.

**Tableau 4**  
***Le nombre d'interventions par entrevue et techniques utilisées par les intervieweurs***

	<b>Nombre d'interventions</b>	<b>Techniques utilisées</b>
Étudiant - Doctorat	22	1. Clarification sur l'item 2. Détection de l'hésitation sur l'item 3. Approfondissement sur l'item 4. Approfondissement sur l'idée et non sur l'item <sup>1</sup> 6. Reformulation de l'item 7. Accompagnement pour donner une réponse
Étudiant - Maîtrise	12	1. Clarification sur le déroulement 2. Renforcement positif 2. Clarification sur le déroulement 3. Clarification sur l'item 4. Accompagnement pour donner une réponse
Professeur 1	27	1. Approfondissement sur l'idée et non sur l'item 2. Paraphrase 3. Accompagnement pour donner une réponse 4. Renforcement positif 1. Clarification sur le déroulement 2. Approfondissement sur l'idée et non sur l'item 3. Détection de l'hésitation sur l'item 4. Accompagnement pour donner une réponse 5. Renforcement positif
Professeur 2	14	1. Clarification sur le déroulement 2. Détection de l'hésitation 3. Reformulation de l'item 4. Clarification sur l'item 5. Approfondissement sur l'idée et non sur l'item 6. Précision sur certains mots

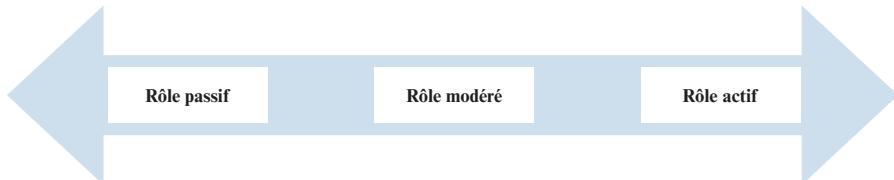
Tableau 4  
*Le nombre d'interventions par entrevue et techniques utilisées  
 par les intervieweurs (suite)*

Nombre d'interventions	Techniques utilisées
24	1. Clarification sur le déroulement 2. Clarification sur l'item 3. Approfondissement sur l'item 4. Reformulation de l'item 5. Renforcement positif 6. Accompagnement pour donner une réponse 7. Détection des hésitations 8. Précision de certains mots
33	1. Clarification sur la démarche 2. Clarification sur l'item 3. Approfondissement sur l'item 4. Approfondissement sur l'idée 5. Détection des hésitations 6. Accompagnement pour donner une réponse
34	1. Clarification sur la démarche 2. Clarification sur l'item 3. Approfondissement sur l'idée 4. Reformulation 5. Accompagnement pour donner une réponse

1.L'interviewé relève le sens du propos et s'éloigne de la réponse à l'item sur lequel il doit se positionner.

À la lumière de ce tableau, nous avons comparé le nombre d'interventions, la variété des techniques et la durée des entrevues pour dégager des profils d'intervieweurs. Pour catégoriser ces différents profils, nous nous sommes appuyées sur les rôles identifiés dans les protocoles à voix haute et de relance. Nos résultats nous ont montré qu'à l'intérieur de ces protocoles, le rôle de l'intervieweur peut varier et s'inscrire sur un continuum de la façon suivante : 1) rôle passif; 2) rôle modéré et 3) rôle actif.

Figure 1

*Illustration des types d'intervieweurs en fonction du nombre, de la variété des interventions et de la durée****Le rôle passif***

Nous avons qualifié le second intervieweur québécois (étudiant à la maîtrise) comme étant passif. Il a réalisé deux entrevues qui se sont déroulées au Québec. La première entrevue a duré 1 heure et 24 minutes, comptait 12 interventions et deux types de techniques; la seconde entrevue a duré 27 minutes, comptait neuf interventions et quatre types de techniques. Nous avons défini ce rôle comme étant passif, puisque l'intervieweur est peu intervenu, et ce, même si la durée des entretiens était longue. De plus, même si les techniques peuvent sembler variées, les types employés sont jugés passifs parce qu'ils portent essentiellement sur la clarification et sur le déroulement ainsi que sur le renforcement positif, mais peu sur le contenu des items et ont peu d'incidence sur la capacité de cerner les problèmes de compréhension spécifiques pour l'interviewé. Des indicateurs comportementaux lors du déroulement reflètent aussi ce rôle passif, comme le manque d'empathie de l'intervieweur qui n'accompagne pas le participant dans une démarche de réponse. Ajoutons aussi le manque d'écoute face à une hésitation de l'interviewé où nous ne relevons pas de réaction ou de position d'écoute de la part de l'intervieweur qui aurait permis au participant de se sentir à l'aise lorsqu'il a émis un doute. Cet intervieweur commence sa maîtrise. Bien qu'il soit détenteur d'un baccalauréat, il a peu d'expérience en recherche, ce qui le conduit à suivre à la lettre le protocole initial de réflexion à voix haute. Il ne s'en écarte pas.

***Le rôle modéré***

Nous avons qualifié de modéré le premier intervieweur, docteurant d'origine africaine et vivant en Ontario ainsi que le professeur 1 d'origine européenne et travaillant en Ontario. Leurs profils sont assez similaires. Le docteurant a réalisé une entrevue en Ontario d'une durée de 54 minutes, a fait 22 interventions et a eu recours à six types de techniques. Le professeur 1,

quant à lui, a réalisé deux entrevues (de 55 minutes et de 49 minutes), toujours en Ontario. Lors de la première entrevue, ce dernier est intervenu 27 fois et a mobilisé quatre types de techniques. Lors de la seconde, ce dernier est intervenu 14 fois et a mobilisé cinq types de techniques. Nous avons qualifié ce rôle de modéré parce que l'intervieweur accompagne davantage le participant dans le processus de réponse que l'intervieweur ayant le premier profil établi. Bien que l'on trouve des relances sur la clarification et sur le déroulement ainsi que du renforcement positif, la différence se manifeste dans la variété des types de techniques qui fait plus écho à l'expression de la pensée du participant, soit par l'empathie relevée durant l'entrevue, soit dans l'accompagnement qui conduit à la réponse. Par exemple, l'intervieweur cherche à clarifier l'idée si cela est nécessaire pour assurer une bonne compréhension à l'interviewé, ou encore il paraphrase ou reformule l'item. En résumé, les interventions se ressemblent sensiblement, même si le professeur a quelques variantes qui s'approchent plus près du participant, par exemple, en relançant l'interviewé. Les caractéristiques de ces deux intervieweurs diffèrent de celles du premier dans la mesure où l'étudiant au doctorat a pris part, avec le chercheur principal, à l'élaboration du questionnaire ainsi qu'à ses modifications. Il était familier avec ce type d'entrevue et avec son contenu. Le second intervieweur, pour sa part, est professeur et a de l'expérience en recherche, ce qui lui donne également de la latitude dans sa façon d'appréhender les items et dans la conduite de l'entrevue.

### *Le rôle actif*

Le professeur 2, d'origine québécoise et vivant en Ontario, a réalisé quatre entrevues en Ontario d'une durée de 58 minutes, 1 heure et 4 minutes, 1 heure et 12 minutes et 1 heure et 8 minutes respectivement, au cours desquelles il est intervenu, dans l'ordre, 23, 24, 33 et 34 fois. Il a fait appel à différents types de relance, dans l'ordre : 6, 8, 6 et 5. Ses types d'interventions sont plus variés que ceux des intervieweurs précédents. Ces données nous amènent à lui attribuer un rôle actif durant l'entrevue. En outre, il est au plus près de l'interviewé, il l'écoute et il l'accompagne. Nous avons également noté des indicateurs comportementaux tels que le fait que l'intervieweur détecte dans les intonations et les longs silences, les hésitations des participants ou encore leurs difficultés de compréhension. Par exemple, la clarification ne porte pas seulement sur l'item. Ici, l'intervieweur a le souci de vouloir bien saisir ce que le participant veut dire et donc, il clarifie aussi l'idée évoquée. De plus, on note un travail d'écoute plus marqué quand il s'agit d'identifier les hésitations de l'interviewé ou

de préciser certains mots. Cet intervieweur est le chercheur principal du projet. Il est familier avec la recherche, avec les deux contextes socioculturels et politiques ainsi qu'avec le devis méthodologique du projet puisqu'il l'a élaboré et adapté. Ces caractéristiques ressortent dans sa manière de se rapprocher du processus cognitif des interviewés et d'adapter ses relances aux problèmes spécifiques de compréhension lors de ces entretiens.

Bien que le nombre, la diversité et la durée des interventions de l'intervieweur ont servi à qualifier les rôles adoptés, il est essentiel d'examiner attentivement la qualité des interactions et des relances. En effet, ce n'est pas uniquement le nombre d'échanges qui permet de comprendre les réponses des participants, mais plutôt la capacité de l'intervieweur à s'ajuster à leurs réactions. Les rôles modéré et actif se démarquent ainsi du rôle passif, car ils favorisent une entrevue plus fidèle à l'expérience et à la pensée de l'interviewé. En somme, l'évaluation du rôle de l'intervieweur ne peut se résumer à des indicateurs quantitatifs, mais doit prendre en compte la dynamique qualitative des échanges entre intervieweurs et interviewés.

### ***La qualité des interventions***

La variation entre ces trois rôles rend compte, d'un point de vue qualitatif, des techniques mobilisées pour faciliter la compréhension des participants quant à leurs interrogations et à leurs hésitations face à un item. C'est pourquoi nous avons analysé également la qualité des interactions entre les intervieweurs et les participants en tenant compte des types de techniques mobilisées. Nous avons catégorisé les extraits de verbatim par type de relances et avons examiné leur qualité en nous attardant sur ce que les intervieweurs ont été en mesure de produire et d'identifier comme problèmes d'interprétation par rapport aux items. Nous avons tenu compte également des conditions de passation en personne (rôle actif) et en virtuel (rôle passif et modéré).

### ***La technique de clarification sur le déroulement ou sur des items***

Dans le tableau 5, nous illustrons à l'aide d'extraits de verbatim, les trois types de rôles adoptés par les intervieweurs et dans quelle mesure la manière dont ils réagissent face aux hésitations des participants favorise l'adaptation des items pour deux échelles : l'attribution causale et les attitudes.

**Tableau 5**  
***La comparaison des profils d'intervieweurs et leur degré d'implication lorsque les participants hésitent à l'égard des items***

Rôle passif	Rôle modéré	Rôle actif
<p><b>Item:</b> Les tests provinciaux sont une très bonne source d'information pour mon enseignement.</p> <p>PS1 : Je pense que c'est une très bonne source d'information dans la mesure où il y a une analyse de tout ça qui est faite. Je m'explique. C'est que si par exemple je suis en 5<sup>e</sup> secondaire, puis on fait vivre l'épreuve unique aux élèves puis que moi, je ne m'en sers pas des résultats. En quoi ça me donne de l'information autre que mes élèves l'ont fait et puis que je sais qu'il y en a tant qui ont réussi. Qu'est-ce qu'on apprend dans le temps? Qu'est-ce qu'on apprend par rapport à l'enseignement? Qu'est-ce qu'on a mis en place dans les dernières années comme formation? Accompagnement? Qu'est-ce qu'on a mis en place comme approche pédagogique novatrice appuyée par la recherche? C'est une excellente source d'information dans la mesure, puis c'est le bout qui me manquerait pour dire que je suis tout à fait d'accord là. Dans la mesure, où on en fait une analyse et qu'on fait quelque chose avec les résultats de cette analyse-là. Si c'est juste de dire, j'ai 28 élèves sur 32 qui ont passé le test. En quoi ça va m'aider dans mon enseignement? Pas grand-chose? Est-ce que je suis tout à fait en accord ou plutôt en désaccord? Il y a entre les deux, il y a une marge-là. Mais ça vous aide pas du tout?</p> <p>I: Oh mais non. Plus on a des commentaires, plus on aime ça.</p> <p>PS1 : Ok, j'en ai moi des commentaires, pas de problème.</p> <p>I: Allez-y, allez-y! Moi, j'ai toute l'après-midi pour vous.</p>	<p><b>Item:</b> Si j'y investis tous mes efforts, j'arrive à aider même les élèves ayant le plus de difficulté ou qui ne sont pas motivés pendant la préparation aux tests provinciaux.</p> <p>PS5 : Si un élève ne peut pas effectuer ce qui est demandé pendant la préparation aux tests provinciaux, je serais capable d'évaluer avec précision si l'apprentissage était à un niveau de difficulté approprié. (longue hésitation)</p> <p>PS5 : Donc si un élève est pas capable d'effectuer ce qui est demandé pendant la préparation je serai capable d'évaluer. Juste avec cette information-là ou est-ce qu'on a...</p> <p>I: Oui tu fais avec l'information qui est là.</p> <p>PS5 : (longue hésitation) Je vais mettre... euh... plutôt en accord.</p> <p>I: Parfait.</p>	<p><b>Item:</b> Les tests provinciaux sont une très bonne source d'information pour mon enseignement.</p> <p>PS10 : Je dirais plus ou moins en accord. Euh j'imagine. Dans ma tête je me justifie, mais je ne l'écrirais peut-être pas. (rires)</p> <p>I: Qu'est-ce qui te ferait hésiter ?</p> <p>PS10: Ben c'est parce que c'est oui puis c'est non en même temps, la réponse à cette question-là... mais expliquer pourquoi ce serait oui puis pourquoi ça serait non ce ne serait pas facile.</p>

Ces extraits montrent trois façons de réagir face aux hésitations des participants concernant un item. L'intervieweur adoptant un rôle passif laisse le participant exprimer longuement son point de vue, ponctué d'interrogations lorsqu'il lui demande si les tests provinciaux sont une très bonne source d'information pour son enseignement. Ayant du mal à se positionner, l'interviewé fait écho à un besoin d'explication de la part de l'intervieweur, mais ce dernier ne détecte pas qu'il y a possiblement une incompréhension ou un problème à clarifier. Pour sa part, l'intervieweur adoptant un rôle modéré est en position d'écoute et réagit face à l'interrogation soulevée par le participant. L'intervieweur confirme son inconfort et l'invite à se positionner sans toutefois demander ce qui peut poser problème. Le participant opère un choix de réponse. L'intervieweur l'invite à nouveau à se positionner. Finalement, l'intervieweur qui joue un rôle actif utilise une technique de reflet face à l'hésitation de l'interviewé, cherchant ainsi à comprendre les éléments qui le perturbent malgré sa réponse. Le participant éprouve de la difficulté à exprimer les réactions que cet item provoque en lui.

#### *La technique de reformulation dans ses propres mots et la paraphrase*

Dans le tableau 6, nous illustrons deux types de rôles adoptés par les intervieweurs et la manière dont les reformulations produites dans leurs propres mots à partir des propos des participants facilitent, ou non, l'adaptation des items pour deux échelles : l'attribution causale et les attitudes.

Dans ces deux extraits, on observe deux façons de reformuler les idées dans les items. L'intervieweur adoptant un rôle modéré intervient à la suite d'une demande de clarification de la part du participant. L'intervieweur reformule en proposant un autre mot pour résoudre ce qui semble un problème lexical dans une phrase courte ou pour atténuer le sens du mot «digne» qui semble déranger l'interviewé. Son emphase sur le mot semble indiquer que le choix de ce terme lui paraît un peu fort. La reformulation sert à cibler le mot problématique. Toutefois, le participant ne semble pas comprendre l'alternative proposée. L'intervieweur tente donc de lui en offrir une autre pour l'aider à comprendre. Le participant n'apprécie pas cette formulation et en demande une autre. Cet échange se termine sur cette demande de reformulation permettant d'identifier le mot problématique sans savoir s'il s'agit d'une préférence ou d'une incompréhension de son sens. L'intervieweur adoptant un rôle actif intervient auprès du

participant lorsque celui-ci semble vivre un conflit cognitif face à une prise de position concernant les facteurs externes influençant les résultats des élèves. L'intervieweur lui propose de reformuler pour l'accompagner dans sa démarche, ce qu'il accepte, et ce qui lui permet de répondre à l'item retenu tout en précisant que la reformulation amène de la clarté, qui est nécessaire pour comprendre l'item.

Tableau 6

***La comparaison des profils d'intervieweurs et de leur reformulation  
dans leurs propres mots des propos du participant***

Rôle modéré	Rôle actif
<p><b>Item: Les tests provinciaux sont dignes de mes efforts.</b></p> <p>PS7: ... sont dignes... Comment vous expliquez digne? Comment ça peut se comprendre?</p> <p>I: Ou alors « méritent que j'y mette des efforts ».</p> <p>PS7: Hum... sont adéquats à... moi j'aurais mis un autre adjectif.... Digne humm... Digne de mes efforts... j'avoue que je ne comprends pas ce que vous voulez dire. Que les tests provinciaux...</p> <p>I: Est-ce que ça vaut la peine que je fasse des efforts?</p> <p>PS7: S'il-vous-plait, essayez de reformuler.</p>	<p><b>Item: Un enseignant est limité dans sa capacité à influencer les résultats des élèves aux tests provinciaux quand tous (accentué) les facteurs sont pris en compte.</b></p> <p>PS10: Tous les facteurs comme humm (silence) Ben de toute façon, même si je ne suis pas sûre, je ne me sens pas limitée. Je dis tout à fait en désaccord de toute façon, mais (soupir) la seule chose. Non c'est pas vrai. Je vais reprendre ça parce que quand tous les facteurs, tu sais c'est juste une fois dans l'année fait que si l'enfant est malade ben je veux dire c'est pas de ma faute à moi et puis c'est une limite, mais ce n'est pas une limite de l'enseignant. C'est ça que j'ai de la misère avec l'énoncé parce que (silence) le résultat est influencé, mais ce n'est pas l'enseignant.</p> <p>I: Puis-je me permettre de reformuler?</p> <p>PS10: Oui.</p> <p>I: Si on disait que quand on parle justement des causes externes, c'est les causes externes à l'enseignant est-ce que ça serait plus clair? Les choses sur lesquelles l'enseignant n'a pas de contrôle.</p> <p>PS10: Oui ce serait beaucoup plus clair parce que là c'est comme si c'était... Je me sens personnellement visée comme si j'avais... Oui reformuler ça aide beaucoup à comprendre.</p> <p>I: Ok!</p> <p>PS10: Donc je répondrais euh sachant cela, je répondrais probablement (relis l'énoncé) ben plutôt en accord à cause de ce facteur. Mais ce serait peut-être bon peut-être d'énoncer certains facteurs, ou même de faire une question à part avec des facteurs plus précis parce que dans le fond on en a quand on parle des capacités des élèves, la performance, les troubles, mais des choses comme ça, mais peut-être ça va venir plus loin, mais c'est ça, comme si je dis l'élève n'est pas à son meilleur une journée ben. Ça c'est ... l'évaluation là, ça arrive.</p>

### ***L'accompagnement du participant pour expliciter sa pensée et donner une réponse***

Dans le tableau 7, nous montrons deux types de rôles adoptés par les intervieweurs et la manière dont leurs relances permettent d'amener les participants à expliciter leur processus mental et à faire des réflexions à voix haute, facilitant ou non l'adaptation des items des deux échelles : celle des attributions causales et celles des attitudes.

Tableau 7

#### ***La comparaison des profils d'intervieweurs et de leur accompagnement vers l'explication du processus mental de réflexion et de choix de la réponse***

Rôle modéré	Rôle actif
<p><b>Item:</b> Les tests provinciaux encouragent les élèves à s'entraider</p> <p>PS5 : (hésitations) En fait, je ne suis pas vraiment en accord avec ça. Je ne pense pas qu'ils encouragent les élèves à s'entraider. Je ne pense pas qu'ils découragent les élèves à s'entraider nécessairement. Mais pas vraiment à s'entraider.</p> <p>I: Donc tu te situerais où là ? Je m'excuse, je te recadre parce qu'il faut vraiment choisir les mesures qui sont là donc sur l'échelle tu prendrais quoi ? Plus ou moins en désaccord ? Plutôt en désaccord ? Ou tout à fait en désaccord ?</p> <p>PS5 : Je pense que je dirais tout à fait en désaccord.</p>	<p><b>Item:</b> Un enseignant est limité dans ce qu'il peut faire étant donné que la motivation et la performance de l'élève aux tests dépend de ses difficultés, comme les difficultés (apprentissage)</p> <p>PS8 : (Longue pause). Ok. Je suis plus ou moins en accord. J'ai besoin de relire beaucoup la question d'avant parce qu'il y a trois actes dans la question : l'effet de l'enseignant, la motivation de l'élève et ses difficultés d'apprentissage. C'est une question qui devient plus complexe vu que ça dépasse deux niveaux de comparaison, mais la question est intéressante. Ouin, plus ou moins en accord.</p> <p>I: Là où ça accrochait le plus dans cette question, comparativement aux autres ?</p> <p>PS8 : Ben je pense que c'est les difficultés d'apprentissage qui jouent parfois souvent sur la motivation de l'élève mais que l'enseignant devrait être habile à fournir à l'élève les stratégies pour surmonter ses difficultés le plus possible. Il a plus un effet, peut-être que je n'ai pas bien répondu. Peut-être que j'ai moins bien répondu, j'aurais dû dire en désaccord parce que l'enseignant est moins limité, il a plus d'impact l'enseignant sur fournir des stratégies à l'élève pour surmonter ces difficultés et l'amener à être plus motivé.</p>

Dans ces extraits, on observe deux façons d'amener les participants à expliciter leur processus de réflexion afin de clarifier une incompréhension de l'item. L'intervieweur adoptant un rôle modéré intervient face à la prise de décision hésitante de l'interviewé. Ce dernier exprime la raison de son hésitation face à l'entraide des élèves lors des tests provinciaux. L'intervieweur amène le participant à répondre en le recadrant et en s'excusant afin de situer sa réponse sur l'échelle. L'intervieweur adoptant

un rôle actif est intervenu après que l'interviewé a pris une longue pause, face à la question qui lui semblait complexe, en l'accompagnant et en lui demandant où se situait le plus ce qui le dérangeait dans l'item. En l'accompagnant dans sa reformulation, l'intervieweur lui a permis d'expliciter sa réflexion et de cibler les termes qu'il faut préciser.

## Discussion et Conclusion

Les constats tirés de cette étude mettent en lumière deux contributions importantes pour la littérature sur l'utilisation des entrevues cognitives pour assurer l'adaptation transculturelle et la validation d'un questionnaire. Dans cet article, nous avons tenté de mettre en avant, dans un premier temps, l'importance du rôle des intervieweurs et de leurs caractéristiques et la manière dont leurs réactions, leurs comportements et leurs relances peuvent influencer la démarche d'adaptation transculturelle et de validation. Bien que des études antérieures aient mis en évidence les apports et les limites des deux protocoles d'entrevue cognitive et les implications pour le rôle des intervieweurs (Beatty & Willis, 2007; Drennan, 2003), cette étude apporte certaines précisions quant aux rôles adoptés par les intervieweurs et les variations qui risquent d'influencer la qualité et la variété des interventions permettant d'adapter et de valider le questionnaire en fonction des contextes socioculturels et politiques distincts.

En comparant les trois types de rôles adoptés par les intervieweurs, nous avons pu dégager certains constats à propos de la variété et de la qualité des techniques ainsi qu'à propos de leur capacité d'extraire des informations verbales. Par exemple, la durée et le nombre d'interventions de l'intervieweur adoptant un rôle passif s'inscrivent dans la démarche du protocole à voix haute. De plus, la qualité et le manque de variété des relances employées portant sur les items spécifiquement ont rendu l'analyse des réponses obtenues difficile à interpréter pour l'adaptation et pour la validation de l'instrument.

Nos résultats montrent également que la reformulation est une relance très spécifique et orientée pour détecter des difficultés dans les items ou dans les structures de phrases (Mohorko & Hlebec, 2016) soulevant des questions importantes quant au moment et à la manière de l'utiliser. La comparaison des techniques qui portent sur la reformulation dans les rôles modéré et actif montre que ce type de relance a une portée limitée

lorsqu'il est utilisé pour des phrases courtes et simples ou encore, par le choix de mots utilisés pour déterminer s'il s'agit d'un problème lexical ou d'un problème de compréhension plus global.

En comparant les trois types de rôles adoptés par les intervieweurs, nous avons pu dégager certains constats liés à leurs comportements, à leurs réactions et à leur capacité à aider le participant à extérioriser sa pensée et à se détendre durant les entretiens (Mohorko & Hlebec, 2016). Nos résultats montrent ainsi que, peu importe le rôle adopté par les intervieweurs, la capacité à s'adapter et à savoir à quels moments intervenir, face à certains comportements ou à certaines réactions (longs silences, hésitations, rires) du participant, est essentielle pour amener ce dernier à extérioriser sa pensée ou pour le sécuriser. Notons également que le renforcement positif, utilisé par les intervieweurs dans les trois types de rôles pour montrer de l'intérêt ou pour valider les propos du participant, semble avoir une portée limitée sur la capacité à sécuriser le participant sur le fait de répondre aux attentes du déroulement de l'entretien.

Cette étude comporte certaines limites liées à notre approche et à la formation des intervieweurs lors de l'adaptation et de la validation d'un questionnaire. Bien que l'étude cherche à comparer la variété et la qualité des techniques utilisées par les intervieweurs ainsi que leur capacité à recueillir des informations verbales, le nombre de participants par province ne nous permet pas de mettre en lumière les aspects socioculturels et politiques distincts qui auraient pu influencer leur compréhension des construits théoriques. Pourtant, le fait de prendre en compte les différences géographiques, politiques et culturelles peut avoir un impact sur l'interprétation des questions et sur les réponses qui sont données, révélant ainsi des sources spécifiques de biais potentiels (Benítez et al., 2018; Lee, 2014).

Cette étude suggère l'importance de former les intervieweurs à mener des entrevues cognitives (Desimone & Le Floch, 2004; Mohorko & Hlebec, 2016) et quelques pistes à considérer dans un contexte d'adaptation transculturelle. Elle montre également que la sensibilisation et la formation des intervieweurs sur les contextes d'application (types d'évaluation, rapports, fonctionnement de l'organisation scolaire) sont essentielles, car les enjeux entourant les examens sont délicats. Cela permettrait de mieux comprendre les informations fournies par les participants et de s'adapter avec souplesse aux besoins et aux contextes spécifiques. Bien qu'une formation initiale ait été mise en place, nous avons observé des limites liées aux consignes

avant et pendant les entrevues ainsi que à l'explication de l'instrument utilisé. Il serait donc pertinent d'envisager une analyse approfondie après chaque entrevue, portant non seulement sur le contenu des propos, mais également sur la qualité des interventions et des réponses obtenues, afin d'accompagner les intervieweurs peu familiers avec cette méthode.

Révision linguistique : Marie-Claire Legaré

Mise en page : Emmanuel Gagnon

Résumé en portugais : Eusébio André Machado

Réception : 14 mai 2024

Version finale : 10 juillet 2024

Acceptation : 18 décembre 2024

## LISTE DES RÉFÉRENCES

- Ajar, D., Dassa, C. & Gougeon, H. (1983). L'échantillonnage et le problème de la validité externe de la recherche en éducation. *Revue des sciences de l'éducation*, 9(1), 3-21. <https://doi.org/10.7202/900396ar>
- Anderson, S. E. & Ben Jaafar, S. (2003). Policy trends in Ontario education. *ICEC Working Paper#1*
- Beatty, P. C. & Willis, G. B. (2007). Research synthesis: the practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287-311. <https://doi.org/10.1093/poq/nfm006>
- Benítez, I., Padilla, J. L., van de Vijver, F. & Cuevas, A. (2018). What Cognitive Interviews Tell Us about Bias in Cross-cultural Research: An Illustration Using Quality-of-life Items. *Field Methods*, 30(4), 277-294.
- Bouchard, S. (2000). Mesurer les variables : quantifier ce qui nous intéresse avec le moins d'erreurs possible. Dans S. Bouchard et C. Cyr (dir.), *Recherche psychosociale : Pour harmoniser recherche et pratique* (p. 231-262). Presses de l'Université du Québec
- Castillo-Díaz, M. et Padilla, J. L. (2013). How cognitive interviewing can provide validity evidence of the response processes to scale items. *Social Indicators Research*, 114, 963-975. <https://doi.org/10.1007/s11205-012-0184-8>
- Copp, D. T. (2016). The impact of teacher attitudes and beliefs about large-scale assessment on the use of provincial data for instructional change. *Education Policy Analysis Archives*, 24, 109. <https://doi.org/10.14507/epaa.24.2522>
- Copp, D. T. (2017). Policy incentives in Canadian large-scale assessment: How policy levers influence teacher decisions about instructional change. *Education Policy Analysis Archives*, 25, 115. <https://doi.org/10.14507/epaa.25.3299>
- Copp, D. T. (2018) Teaching to the test: a mixed methods study of instructional change from large-scale testing in Canadian schools. *Assessment in Education: Principles, Policy & Practice*, 25(5), 468-487. <https://eric.ed.gov/?id=EJ1195192>
- Corbière, M. & Fraccaroli, F. (2020). La conception, la validation, la traduction et l'adaptation transculturelle d'outils de mesure : des exemples en santé mentale et travail. Des exemples en santé mentale et travail. Dans M. Corbière et N. Larivière (dir.), *Méthodes qualitatives, quantitatives et mixtes*, 2<sup>e</sup> édition : Dans la recherche en sciences humaines, sociales et de la santé (p. 703-752). Presses de l'Université du Québec. <https://doi.org/10.2307/j.ctv1c29qz7.32>
- Desimone, L. M. & Le Floch, K. C. (2004). Are we asking the right questions ? Using cognitive interviews to improve surveys in education research. *Educational Evaluation and Policy Analysis*, 26(1), 1-22. <https://doi.org/10.3102/01623737026001001>
- Drennan, J. (2003). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing*, 42(1), 57-63. <https://doi.org/10.1046/j.1365-2648.2003.02579.x>
- Dunn, R., Jaafar, S. B., Earl, L. & Katz, S. (2013). Towards data-informed decisions: from ministry policy to school practice. *Data-based Decision Making in Education: Challenges and Opportunities*, 155-175. [https://link.springer.com/chapter/10.1007/978-94-007-4816-3\\_9](https://link.springer.com/chapter/10.1007/978-94-007-4816-3_9)

- Fullan, M. (2000). The return of large-scale educational reform. *Journal of Educational Change*, 1(1), 5-28.
- Gagnon, M. P., Attieh, R., Dunn, S., Grandes, G., Bully, P., Estabrooks, C. A. & Ouimet, M. (2018). Development and content validation of a transcultural instrument to assess organizational readiness for knowledge translation in healthcare organizations: the OR4KT. *International Journal of Health Policy and Management*, 7(9), 791. <https://doi.org/10.15171/ijhpm.2018.17>
- Gana, K., Boudouda, N. E., Ben Youssef, S., Calcagni, N. & Broc, G. (2021). Adaptation transculturelle de tests et échelles de mesure psychologiques : Guide pratique basé sur les Recommandations de la Commission Internationale des Tests et les Standards de pratique du testing de l'APA. *Pratiques psychologiques*, 27(3), 223240. <https://doi.org/10.1016/j.prps.2021.02.001>
- García, A. A. (2011). Cognitive Interviews to Test and Refine Questionnaires. *Public Health Nursing*, 28(5), 444-450. <https://doi.org/10.1111/j.1525-1446.2010.00938.x>
- Grondin, J., Dionne, É., Fleuret, C. & Boiteau, N. (2015). Sources d'invalidité et d'erreur dans la traduction ou l'adaptation de tests: un état de la question. *Revue de l'Université de Moncton*, 46(1-2), 291-323. <https://doi.org/10.7202/1039041ar>
- Gudmundsson, E. (2009). Guidelines for translating and adapting psychological instruments. *Nordic Psychology*, 61(2), 29-45. <https://doi.org/10.1027/1901-2276.61.2.29>
- Lauzier, M., Côté, K., Annabi, D. & Melançon, S. (2021). La validation transculturelle d'instruments de mesure en psychologie : Un portrait des pratiques utilisées dans les travaux publiés entre 1989 et 2019. [Cross-cultural validation of measuring instruments in psychology: A portrait of the practices used in the works published between 1989 and 2019.]. *Canadian Psychology/Psychologie canadienne*, 117. <https://doi.org/10.1037/cap0000302>
- Lee, J. (2014). Conducting cognitive interviews in cross-national settings. *Assessment*, 21(2), 227-240.
- Loye, N. (2018). Et si la validation était plus qu'une suite de procédures techniques ? *Mesure et Évaluation en Éducation*, 41(1), 97-123. <https://doi.org/10.7202/1055898ar>
- Maroy, C., Mathou, C. & Vaillancourt, S. (2017). Les enseignants québécois à l'épreuve de la «gestion axée sur les résultats». Une étude qualitative dans quatre établissements secondaires. *Dans, RASE – Revista de la Asociacion de Sociologia de la Educacion*, 10(3), 539-557. <http://dx.doi.org/10.7203/RASE.10.3.9906>
- Ministère de l'Éducation de l'Ontario (2022). *Écoles de langue française*. <https://www.ontario.ca/fr/page/leducation-en-langue-francaise>
- Ministère de l'Éducation et de l'Enseignement supérieur. (2015). *Statistiques de l'éducation Éducation préscolaire, enseignement primaire et secondaire*. [https://www.education.gouv.qc.ca/fileadmin/site\\_web/documents/PSG/statistiques\\_info\\_decisionnelle/15-00503\\_statistiques\\_2015\\_edition\\_v25oct.pdf](https://www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/statistiques_info_decisionnelle/15-00503_statistiques_2015_edition_v25oct.pdf)
- Mohorko, A. & Hlebec, V. (2016). Degree of cognitive interviewer involvement in questionnaire pretesting on trending survey modes. *Computers in Human Behavior*, 62, 79-89.
- Priede, C. & Farrall, S. (2011). Comparing results from different styles of cognitive interviewing: 'verbal probing' vs. 'thinking aloud'. *International Journal of Social Research Methodology*, 14(4), 271-287.

- Priede, C., Jokinen, A., Ruuskanen, E. & Farrall, S. (2014). Which probes are most useful when undertaking cognitive interviews? *International Journal of Social Research Methodology*, 17(5), 559–568. <https://doi.org/10.1080/13645579.2013.799795>
- Ridolfo, H. & Schoua-Glusberg, A. (2011). Analyzing cognitive interview data using the constant comparative method of analysis to understand cross-cultural patterns in survey data. *Field Methods*, 23(4), 420-438.
- Rozenwajn, E. & Dumay, X. (2016). Les évaluations externes à faibles enjeux : les enseignants confrontés à eux-mêmes ? *Revue française de pédagogie*, (194), 71-90. <https://doi.org/10.4000/rfp.4977>
- Sabourin, S., Valois, P. & Lussier, Y. (2000). L'utilisation des questionnaires en recherche : une solution pratique qui nécessite une démarche rigoureuse. Dans S. Bouchard et C. Cyr (dir.), *Recherche psychosociale : pour harmoniser recherche et pratique* (p. 263-304), Presses de l'Université du Québec.
- Sopromadze, N. & Moorosi, P. (2017). Do we see through their eyes? Testing a bilingual questionnaire in education research using cognitive interviews. *International Journal of Research & Method in Education*, 40(5), 524-540.
- Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly*, 79(S1), 359-395.
- Willis, G. B., & Miller, K. (2011). Cross-cultural cognitive interviewing: Seeking comparability and enhancing understanding. *Field methods*, 23(4), 331-341.
- Winton, S. & Milani, M. (2017). Policy advocacy, inequity, and school fees and fundraising in Ontario, Canada. *Education Policy Analysis Archives*, 25, 40-40. <https://doi.org/10.14507/epaa.25.2679>
- Yerly, G. (2017). Les raisons du faible usage des résultats d'évaluation externe par les enseignants. Étude croisée dans trois contextes éducatifs. *Contextes et didactiques*, (9), 60-71. <https://doi.org/10.4000/ced.745>