

Université de Montréal

Analyse de la qualité méthodologique des outils de mesure du leadership tyrannique :
les chercheurs et les praticiens possèdent-ils des instruments de mesure adaptés
au système scolaire québécois?

Par

Joelle Veilleux Deschênes

Département de psychologie, Faculté des arts et des sciences

Essai présenté en vue de l'obtention du doctorat (D. Psy.)
en psychologie, option psychologie du travail et des organisations

Mars 2021

© Joelle Veilleux Deschênes, 2021

Université de Montréal

Département de psychologie, Faculté des arts et des sciences

Cet essai est intitulé

**Analyse de la qualité méthodologique des outils de mesure du leadership tyrannique :
les chercheurs et les praticiens possèdent-ils des instruments de mesure adaptés
au système scolaire québécois?**

Présenté par

Joelle Veilleux Deschênes

A été évalué par un jury composé des personnes suivantes

Jean-Sébastien Boudrias

Président-rapporteur

Carole Sénéchal

Directrice de recherche

Luc Brunet

Membre du jury

Résumé

Le leadership tyrannique est un type de leadership destructeur où le leader peut utiliser l'humiliation, le dénigrement et la manipulation des employés afin de les influencer à atteindre des objectifs de travail. Plus d'un travailleur sur dix en serait victime aux États-Unis. Ce type de leader peut occasionner des conséquences dévastatrices sur les employés, les équipes de travail et les organisations. Une meilleure compréhension des antécédents et des mécanismes qui promeuvent le leadership tyrannique au sein des organisations permettrait d'adapter les processus de sélection et de formation afin de réduire l'occurrence de ce phénomène ainsi que les conséquences néfastes qu'il engendre. Il existe un nombre limité d'études réalisées sur ce sujet dans le secteur de l'éducation, alors que plusieurs variables associées au leadership tyrannique sont actuellement observables dans le quotidien des administrateurs scolaires et des enseignants (ex : demandes élevées liées à l'emploi, manque de ressources, haut niveau d'épuisement professionnel et de détresse psychologique), soulignant le besoin d'étudier le phénomène au sein de ce milieu de travail. Toutefois, il impose d'avoir des outils qui permettent de bien capter le construit. Cela dit, plusieurs enjeux liés à la mesure du leadership tyrannique nécessitent une attention particulière, notamment la multiplicité des outils utilisés et leurs variantes, le non-consensus quant à la dimensionnalité, le recours à une diversité de sources et la variabilité individuelle ou culturelle possible quant à ce qui est perçu comme de l'abus. Une meilleure compréhension de ceux-ci permettrait d'effectuer des choix méthodologiques plus éclairés pour contrer les limites des outils actuels. Afin de répondre à ce besoin, l'objectif de la présente étude consiste à effectuer une revue systématique des outils de mesure du leadership tyrannique et de discuter de leur qualité en fonction des enjeux de mesure préalablement ciblés, de sorte à identifier comment une mesure validée pourrait être adaptée au milieu scolaire québécois. Le but ultime est d'aider les chercheurs et les praticiens dans le choix et l'adaptation, si nécessaire, de leur instrument de mesure afin de capter le construit adéquatement.

Mots-clés : leadership tyrannique, supervision abusive, milieu scolaire, mesure, artéfacts de mesure, psychologie du travail et des organisations

Abstract

Tyrannical leadership is a type of destructive leadership in which the leader uses humiliation, belittlement and manipulation to influence employees in achieving work-related goals. In the United States more than one in ten workers is a victim of this type of leadership. Such a leader can have devastating consequences for employees, work teams and organisations. A better understanding of the antecedents and mechanisms that promote tyrannical leadership within organisations would allow the tailoring of selection and training processes in order to reduce its occurrence and the negative consequences it engenders. There are a limited number of studies that have been carried out in the education sector. Several variables associated with the construct are observable in the daily life of school administrators and teachers (for example, high work demands, lack of resources, high level of burnout and psychological distress), highlighting the need to study tyrannical leadership in this environment. However, this requires having tools that allow us to properly capture the construct. Several issues related to the measurement of tyrannical leadership require particular attention, such as the multiplicity of tools and their variants, the non-consensus with regards to the dimensionality, the use of a diversity of sources and the possible individual or cultural variability as to what is perceived as abuse. A better understanding of these would shed light on methodological choices that could counter the limitations of current tools. In order to meet this need, the objective of this study is to carry out a systematic review of tools used to measure tyrannical leadership and discuss their quality according to the targeted measurement issues, so as to identify how a validated measure could be adapted to the Quebec school environment. The goal is to help researchers and practitioners in choosing and adapting, if necessary, their measuring instrument in order to more effectively capture the construct.

Keywords: tyrannical leadership, abusive supervision, education, measure, measurement artefacts, work and organisational psychology

Table des matières

Résumé.....	3
Abstract.....	4
Table des matières.....	5
Liste des tableaux.....	8
Liste des figures.....	9
Liste des sigles et abréviations.....	10
Remerciements.....	12
Contexte théorique.....	13
Position du problème.....	13
Leadership tyrannique.....	14
La mesure et ses enjeux.....	16
Multiplicité de versions.....	16
Multiplicité des conceptualisations.....	17
Diversité de sources.....	18
Récurrence et prévalence.....	18
Exhaustivité et perception de victimisation.....	19
Comparaisons organisationnelles et ethniques.....	20
Le secteur de l'éducation.....	21
Méthode.....	23
Sélection des études.....	23
Qualité méthodologique des outils.....	24
Résultats.....	25
Description des outils recensés.....	27

Outil de Tepper (2000) et ses variantes.....	28
Johnson et al. (2012).....	29
Ashforth (1994).....	30
Cacciatore (2015).....	30
Schmidt (2008).....	30
Asland et al. (2010).....	31
Larsson et al. (2012).....	31
Thoroughgood et al. (2012).....	31
Résultats obtenus aux objectifs de recherche.....	32
Multiplicité des versions de l’outil de Tepper (2000).....	32
Multiplicité des conceptualisations.....	35
Diversité de sources.....	38
Supervision abusive vicariante.....	38
Récurrence et prévalence.....	40
Récurrence et devis transversaux.....	40
Récurrence et devis longitudinaux.....	40
Récurrence et prévalence.....	40
Exhaustivité et perception de victimisation.....	41
Comparaisons organisationnelles et ethniques.....	42
Comparaisons ethniques.....	45
Comparaisons organisationnelles.....	46
Discussion.....	47
Constats, pistes de réflexion et recommandations sur les outils à partir des enjeux identifiés.....	47
Multiplicité de versions.....	47

Multiplicité des conceptualisations.....	49
Diversité de sources.....	49
Récurrence et prévalence.....	50
Exhaustivité et perception de victimisation.....	51
Comparaisons organisationnelles et ethniques.....	53
Somme toute.....	54
Limites de l'étude.....	55
Conclusion.....	55
Références.....	57
Annexes.....	70

Liste des tableaux

Tableau 1. <i>Liste des échelles incluses dans le cadre du présent essai, le construit qu'elles mesurent et le nombre d'études recensées qui en font l'utilisation.....</i>	27
Tableau 2. <i>Compilation d'informations extraites à l'égard des outils de mesure recensés.....</i>	27
Tableau 3. <i>Artéfacts de mesure de la perception de supervision abusive du présent essai.....</i>	33
Tableau 4. <i>Artéfacts de mesure de la perception de supervision abusive obtenus par Mackey et al. (2017).....</i>	33
Tableau 5. <i>Résultats obtenus avec l'outil de Tepper (2000) par pays.....</i>	43
Tableau 6. <i>Comparaison des items de l'outil de Tepper (2000) avec les manifestations abusives les plus fréquentes ainsi que celles occasionnant la plus grande intensité de souffrance identifiées par une population d'enseignants dans l'article de Blase, Blase et Du (2008).....</i>	51
Annexe A	
Annexe A. <i>Explications des choix d'exclusion des construits.....</i>	70
Annexe B	
Tableau B1. <i>Coefficient de corrélation (r) entre la variable dichotomique regroupant les études ayant indiqué le type d'échelle utilisé en un groupe et ceux ne l'ayant pas indiqué en un autre groupe.....</i>	73
Tableau B2. <i>Coefficient de corrélation (r) entre la variable dichotomique regroupant les études ayant utilisé cinq ou quinze items en un groupe et celles ayant utilisé un autre nombre d'items en un autre groupe.....</i>	74
Tableau B3. <i>Statistiques descriptives de l'ANOVA réalisée afin de répondre à l'objectif 1.....</i>	74
Tableau B4. <i>Statistiques descriptives de l'ANOVA réalisée afin de répondre à l'objectif 6.....</i>	74

Liste des figures

Figure 1. <i>Processus de sélection des articles</i>	26
Figure 2. <i>Moyenne de supervision abusive selon le nombre d'items de l'outil de Tepper (2000) et le type d'échelle utilisés</i>	35
Figure 3. <i>Moyenne de supervision abusive obtenue par l'outil de Tepper (2000) selon le type d'échelle utilisé ainsi que l'appartenance à une région du globe (Orient c. Occident)</i>	45

Liste des sigles et abréviations

SA : supervision abusive

ANOVA : analyse de variance (*analysis of variance*)

À ma famille

À ma directrice de recherche

Remerciements

Avec humilité, je tiens à souligner à quel point plusieurs personnes ont contribué de façon plus ou moins directe à rendre cet essai possible.

Jacob, mon grand garçon et mon plus grand amour... tu es ma source principale de motivation. Toutes mes décisions et mes actions, incluant les efforts que j'ai investis dans le cadre du présent essai, sont faites en ayant ton bien-être à cœur. Tu fais ressortir le meilleur de moi-même.

Daniel, Thérèse et Denis... vous êtes mes piliers. Vous êtes présents pour moi depuis toujours et je ne serais tout simplement pas là où je suis sans vous.

Ma famille... au sens large du terme. Vous êtes mes amis chers, ceux qui partagez mes joies et mes peines. Vous contribuez à rendre ma vie si belle.

Mme Sénéchal... vous avez su trouver les bons mots à dire aux bons moments. Vous êtes accessible et présente comme nulle autre pareille. Vous avez rendu l'exercice que consiste cet essai fort agréable, contribuant ainsi à faire de mon parcours universitaire une si belle aventure.

M. Larivée... merci pour avoir généreusement lu et relu mes travaux. Merci aussi pour votre œil aiguisé à repérer les améliorations possibles ainsi que pour votre rétroaction tellement appréciée.

M. Boudrias et M. Brunet... merci pour l'intérêt que vous portez à mes travaux et pour vos commentaires si précieux. Vous avez à cœur le bien de vos étudiants. Je suis honorée que vous soyez sur mon jury d'évaluation.

Mes chers collègues de l'Université de Montréal, ceux et celles qui ont été à mes côtés tout au long du doctorat... J'ai ris et/ou pleuré avec plusieurs d'entre vous. Les souvenirs de mon parcours doctoral restent gravés en moi pour toujours. Merci d'en avoir fait partie.

« Vous savez, on peut trouver du bonheur même dans les endroits les plus sombres.

Il suffit de se souvenir d'allumer la lumière »

(Dumbledore, Harry Potter et le prisonnier d'Azkaban)

Contexte théorique

Position du problème

Le leadership tyrannique et la supervision abusive, considérés ici comme conceptuellement synonymes (Einarsen et al., 2007; Krasikova et al., 2013), ont fait l'objet d'un nombre grandissant d'études au cours des dernières décennies (Ashforth, 1994; Aasland et al., 2010; Tepper et al., 2017). Un leader tyrannique poursuit les buts visés par l'organisation au dépend de la motivation, du bien-être ou de la satisfaction des employés (Einarsen et al., 2007). Autrement dit, de tels leaders « peuvent humilier, dénigrer et manipuler les employés pour que le travail se fasse » (traduction libre, Aasland et al., 2010, p. 441). Aux États-Unis, 13,6 % des travailleurs seraient victimes de supervision abusive (Tepper, 2007), ce qui est associé à un coût annuel de 23,8 milliards de dollars américains (dû par exemple à l'absentéisme, au taux de roulement et à la baisse de productivité qui en résultent; Tepper et al., 2006), quoique la validité de ce constat soit controversée (Martinko et al., 2013).

Les impacts dévastateurs de ce phénomène sur les individus et les organisations ont motivé les chercheurs et les praticiens du domaine à vouloir mieux discerner sa présence et identifier des stratégies permettant d'en réduire l'occurrence. Cet engouement a mené à la publication de maintes études et méta-analyses, notamment sur les antécédents, les conséquences et les corrélats du leadership tyrannique (ex. : Mackey et al., 2017; Tepper et al., 2017; Zhang et Bednall, 2016). Malgré tout, de multiples questions de recherche persistent et la plupart des publications se concentrent sur les conséquences du phénomène (Tepper et al., 2017). L'urgence de développer des modèles théoriques intégratifs plus complets et de mieux cerner les antécédents se fait ressentir (Zhang et Bednall, 2016). D'un point de vue pratique, une meilleure compréhension des antécédents et des mécanismes qui promeuvent le leadership tyrannique au sein des organisations pourrait influencer les processus de sélection et de formation, de manière à réduire l'occurrence et l'intensité de ce type de leadership. Un nombre encore plus limité d'études a été réalisé dans le secteur de l'éducation (Blase, Blase et Du, 2008). Il s'agit pourtant d'un contexte de travail possédant des caractéristiques qui favorisent la perception de supervision abusive (Tepper et al., 2017), telles qu'une surcharge de travail, des demandes élevées liées à l'emploi et un manque de ressources.

Cela dit, plusieurs auteurs ont identifié des enjeux liés aux outils de mesure et à la façon de les utiliser susceptibles d'entraîner des conclusions erronées, et ainsi nuire à la compréhension du phénomène (ex. : Mackey et al., 2017; Sasso et Gonzalez-Morales, 2018). Notamment, on remarque une multiplicité de façons dont les outils ont été adaptés; par exemple, certains ont recours à un nombre réduit d'items ou à divers types d'échelle (Cacciatore, 2015). Les instruments divergent également quant au nombre de dimensions que comporte le construit (ex. : Mitchell et Ambrose, 2007). Mackey et ses collègues (2017)

soulignent d'ailleurs la possibilité que la dimensionnalité du construit puisse être différente tout dépendant du type d'organisation étudié (c'est-à-dire que la supervision abusive pourrait se manifester différemment selon le contexte). En outre, selon le contexte organisationnel et ethnique étudié, les mesures comportementales existantes pourraient ne pas être exhaustives ou, à l'inverse, certains items que contiennent les outils pourraient ne pas être considérés comme abusifs par les répondants, présumés victimes (Hershcovis, 2011; Sasso et Gonzalez-Morales, 2018).

Néanmoins, à ce jour, peu d'études portent une attention particulière sur les enjeux de mesure et leur influence sur la recherche et la pratique. Une meilleure compréhension des limites méthodologiques des outils de mesure nous permettrait de nuancer l'interprétation des données que nous faisons et ainsi accroître notre confiance à l'égard des connaissances générées. C'est aussi en ayant une meilleure connaissance des limites de nos outils que nous pouvons faire des choix méthodologiques afin de mieux les contrer. Par ailleurs, le milieu scolaire possédant des caractéristiques qui lui sont propres, il importe de s'assurer qu'une échelle de mesure du leadership tyrannique soit validée en contexte scolaire avant qu'elle ne soit utilisée par des chercheurs ou des professionnels.

Plus concrètement, le but du présent article consiste à effectuer une revue systématique de la littérature afin de repérer l'ensemble des mesures du leadership tyrannique utilisées en milieu organisationnel au cours des cinq dernières années et d'extraire les informations ayant trait à divers enjeux de mesure. Par la suite, un bilan et une analyse critique de la qualité méthodologique de ces outils tenteront de déterminer comment un outil de mesure pourrait être adapté et validé afin d'être utilisé au sein du système scolaire québécois.

Leadership tyrannique

Ashforth (1994) est le premier à utiliser les termes *petty tyrant* qu'il définit comme « celui qui exerce son pouvoir sur les autres » (traduction libre, p. 755). Ashforth donne des exemples de comportements associés à un leader tyrannique, ceux-ci ne constituant pas une liste exhaustive : agit avec méfiance et suspicion, supervise étroitement ses subordonnés, critique les employés de manière dure et en public, agit de façon condescendante, s'attribue le mérite du travail des autres, blâme autrui pour ses erreurs, peut avoir des sautes d'humeur, ne consulte pas, nuit au développement d'autrui, a tendance à surcontrôler ses subordonnés et à les traiter de manière arbitraire, punitive et sans égard à leur bien-être.

Un leader tyrannique peut arborer des caractéristiques qui paraissent positives aux yeux des cadres supérieurs et des employés d'une organisation; c'est le cas par exemple du charisme, de l'attitude visionnaire et de l'extraversion qui aident le leader narcissique à faire une bonne première impression (Braun, 2017; Campbell et Campbell, 2009). De plus, étant donné que le leadership tyrannique peut

permettre d'atteindre certains buts organisationnels (Einarsen et al., 2007), le leader peut être vu d'un très bon œil par la haute direction tout en étant perçu comme très toxique/intimidateur par les employés. Néanmoins, les effets ou les attributs positifs qui pourraient être associés à un leader abusif (ex. : émerger en tant que leader; paraître confiant, dominant et/ou performant) seraient de courte durée (moins de douze semaines selon Ong et al., 2016, après quoi les effets négatifs se font ressentir), puisque la qualité des relations d'un tel leader s'érode au fil du temps.

Similairement, Aasland et ses collaborateurs (2010) suggèrent que, si les comportements tyranniques ne sont pas traités, le leader tyrannique pourrait « dérailler » (*derailed leadership*), impliquant que ses comportements nuiraient éventuellement non seulement aux subordonnés, mais également à l'organisation. Autrement dit, les effets positifs qu'un tel leader aurait sur les buts de l'organisation pourraient disparaître (Ma et al., 2004) et, à titre d'exemple, engendrer une baisse de rendement ou une augmentation du taux de roulement (Tepper et al., 2006). En ce sens, l'étude du leadership tyrannique pourrait permettre d'identifier les leaders abusifs avant qu'ils ne « déraillent ».

Ashforth (1994) présente un outil de mesure de quarante-sept items qui regroupe le construit en six dimensions : (1) faire preuve de gestion arbitraire; (2) déprécier ses subordonnés; (3) manquer de considération; (4) décourager l'initiative de ses subordonnés; (5) appliquer des punitions non contingentes; et (6) avoir un style de résolution de conflit autocratique (étiquettes françaises des dimensions rapportées par Forget, 2016). Cacciatore (2015) explicite certaines précisions à l'égard du leadership tyrannique d'Ashforth (1994), soit que : (1) les six dimensions de son outil comportent à la fois des comportements hostiles (ex : critiquer publiquement ses subordonnés) et des comportements obstructifs (ex : ne consulte pas autrui ou ne les tient pas informés; (2) les leaders tyranniques « peuvent ne pas être conscients de leur comportement et de ses effets » (traduction libre, Ashforth, 1994, p. 772), c'est-à-dire qu'un leader tyrannique choisit ce type de comportements plutôt que des comportements constructifs et peut justifier l'adoption de tels comportements par des raisons qui lui semblent valables; et (3) les comportements physiques sont exclus de la définition du construit. Même si les actes physiques peuvent être dévastateurs, Cacciatore (2015) explique qu'ils ne persistent généralement pas longtemps au sein des organisations; les comportements destructeurs étant plutôt insidieux, c'est-à-dire qu'ils s'installent graduellement et affaiblissent la qualité des relations au fil du temps.

Plusieurs auteurs considèrent la supervision abusive de Tepper (2000) comme un construit conceptuellement comparable au leadership tyrannique (Cacciatore, 2015; Einarsen et al., 2007; Krasikova et al., 2013; Schyns et Schilling, 2013; Tepper, 2007; Zellars et al., 2002), et certains utilisent même les termes de façon interchangeable. Tepper (2000) définit la supervision abusive par « la perception des subordonnés quant à quel point les superviseurs adoptent de manière soutenue des

comportements hostiles, verbaux ou non verbaux, en excluant les actes physiques » (traduction libre, p. 178). D'un point de vue empirique, les deux construits seraient hautement corrélés ($r = 0,77$; Boudrias et al., 2020; Cacciatore, 2015). Dans une analyse de profils latents effectuée à partir de résultats obtenus à l'aide d'échelles mesurant dix types de leadership différents, Gatti et al. (2019) ont trouvé que les données de supervision abusive (mesurée par l'outil de Tepper, 2000) ainsi que celles de la *Petty Tyranny* (mesurée par neuf items de l'outil d'Ashforth, 1994) montraient un pattern très similaire entre les six profils de leaders identifiés. Ils en ont conclu qu'ils devraient être traités comme un seul construit. Ainsi, étant donné que les définitions sont très semblables, et qu'un certain soutien empirique supporte cette similarité conceptuelle (ex. : Einarsen et al., 2007; Krasikova et al., 2013; Gatti et al., 2019), l'auteure du présent texte s'intéresse aux outils de mesure du leadership tyrannique et de la supervision abusive. Elle utilise les deux appellations de façon interchangeable.

La mesure et ses enjeux

Le présent essai se concentre sur les six enjeux de mesure décrits ci-dessous. Ceux-ci mettent en lumière certaines limites potentielles des outils de mesure. Pour chacun des enjeux, des objectifs de recherche précis sont présentés; ceux-ci guideront ultérieurement le bilan et l'analyse qui seront effectués concernant les outils repérés par la recension des écrits.

Multiplicité de versions. Dans les écrits empiriques, l'échelle de Tepper (2000) est de loin l'outil le plus utilisé pour mesurer la supervision abusive (Mackey et al., 2017; Tepper, 2007). Celui-ci a été traduit dans plus d'une quinzaine de langues (voir Lavoie-Tremblay et al., 2016 pour la version canadienne-française) et appliqué aux quatre coins du globe (Martinko et al., 2013). L'outil original de Tepper (2000) possède quinze items et mesure le construit à l'aide d'une échelle de fréquence comportant cinq points d'ancrage (1 = *I cannot remember him/her ever using this behavior with me* à 5 = *He/she uses this behavior with me very often*).

Cela dit, une variété de versions de cet outil a été répertoriée par Mackey et al. (2017). Dans leur méta-analyse, qui recensait les articles publiés avant 2014 inclusivement, les auteurs rapportent que certaines études utilisent des versions comportant de 3 à 19 items; certaines utilisent une échelle d'accord (ex : fortement en désaccord à fortement en accord) plutôt qu'une échelle de fréquence; et le nombre de points d'ancrage varie beaucoup (allant de 0 à 10; de 1 à 5 ou de 1 à 7, tout type d'échelle confondu). Dans ce même ordre d'idées, Cacciatore (2015) souligne la nécessité « d'unifier les instruments de mesure » (traduction libre, p. 33) ainsi que la tendance des chercheurs à préférer l'utilisation de questionnaires courts.

En ce sens, Mackey et al. (2017) ont examiné les artéfacts de mesure liés aux multiples adaptations de l'instrument de Tepper (2000). En ce qui concerne le type d'échelle utilisé (codé comme suit par les auteurs pour les calculs statistiques : fréquence = 0; accord = 1), celui-ci était modérément corrélé aux moyennes ($r = 0,29$), aux écart-types ($r = 0,28$) et aux indices de fidélité (α ; $r = 0,29$) rapportés dans les écrits (quoique les indices de fidélité étaient élevés pour la plupart des études, $\mu = 0,92$, $\sigma = 0,04$). La direction de la corrélation montre que les moyennes de supervision abusive obtenues à l'aide d'une échelle d'accord auraient tendance à être plus élevées que celles obtenues par une échelle de fréquence. Ces constats indiquent que le choix du type d'échelle pourrait avoir un impact sur les résultats obtenus dans les études empiriques et doit donc être étudié davantage. En revanche, le nombre d'items utilisés était peu corrélé aux moyennes ($r = -0,03$), aux écart-types ($r = -0,07$) et aux indices de fidélité (α ; $r = -0,09$). Similairement, Breevhart et de Vries (2017) ont obtenu une corrélation de l'ordre de la singularité ($r = 0,90$) entre la version originale de Tepper (2000) à quinze items et celle de Mitchell et Ambrose (2007) à cinq items. De prime abord, les résultats semblent indiquer qu'il serait acceptable d'utiliser une version plus courte de l'outil, mais il s'impose d'examiner ces corrélations plus en profondeur avant de pouvoir réellement affirmer une telle conclusion.

Objectif 1 : reproduire les analyses portant sur les artéfacts de mesure effectuées par Mackey et al. (2017) en utilisant les articles publiés entre 2015 et 2020, afin de vérifier si le type d'échelle et le nombre d'items utilisés de l'outil de Tepper (2000) sont corrélés à la moyenne et à l'écart-type (obtenus pour la supervision abusive) ainsi qu'à l'indice de fidélité rapportés dans les écrits.

Multiplicité des conceptualisations. Le nombre de dimensions que comporte le construit est une autre caractéristique de l'outil qui ne fait pas consensus. Tepper (2000) avait initialement trouvé une seule dimension, alors que Mitchell et Ambrose (2007) en ont trouvé deux : les comportements abusifs passifs-agressifs et les comportements abusifs actifs-agressifs. De surcroît, en 2012, Tepper, Duffy et Breaux-Soignet ont décrit deux types de supervision abusive : le type impulsif (actions automatiques et non contrôlées) et le type stratégique (actions délibérées ayant un but particulier). Cela dit, il importe de clarifier en quoi les diverses façons de concevoir la dimensionalité du construit sont liées entre elles et de vérifier en quoi elles possèdent une validité incrémentielle.

Objectif 2 : parmi les écrits publiés entre 2015 et 2020, identifier quelles sont les diverses façons de concevoir la dimensionnalité du construit et repérer les articles qui examinent empiriquement en quoi cette diversité de conceptualisations permet de mieux prédire (ou de prédire différemment) le réseau nomologique de la variable.

Diversité de sources. Les questionnaires mesurant le leadership tyrannique sont habituellement complétés par les subordonnés qui se prononcent sur les comportements de leur superviseur (ex. : Tepper, 2000), adoptant ainsi une approche centrée sur la victime. Plusieurs études récentes ont néanmoins eu recours à des données auto-rapportées par les leaders afin de démontrer un soutien empirique aux modèles testés (ex. : Courtright et al., 2016; Johnson et al., 2012; Liao et al., 2018), optant pour une approche centrée sur le leader. Certains choisissent aussi d'administrer le questionnaire aux collègues de travail, que ce soit pour obtenir une mesure plus objective de la supervision abusive provenant d'une tierce partie (Lukacik et Bourdage, 2019) ou parce qu'ils s'intéressent à la supervision abusive vicariante et à ses effets (ex. : Kluemper et al., 2019). Néanmoins, en quoi les données provenant de diverses sources (ex : leader, subordonnés, collègues) sont liées les unes aux autres et les facteurs influençant ces corrélations demeurent peu étudiés (Ambrose et Ganegoda, 2020; Martinko et al., 2013; Tepper et al., 2017).

Objectif 3 : parmi les articles publiés au cours des cinq dernières années, décrire les tendances quant à qui (leader, subordonnés, collègues) complètent les outils de mesure et repérer les études qui comparent empiriquement les données de supervision abusive obtenues de diverses sources.

Récurrence et prévalence. Souvent, les outils sont utilisés pour mesurer le construit à un seul moment dans le temps (Mackey et al., 2017), ou on demande aux participants de répondre aux items en réfléchissant aux comportements de leur superviseur survenus antérieurement dans un certain intervalle de temps (ex. : Dionisi et Barling, 2019). Ceci ne permet pas nécessairement de capter l'aspect de récurrence des comportements abusifs, tel qu'il est compris dans les définitions du construit, ou ne permet pas d'analyser les nuances ni l'évolution des manifestations comportementales au fil du temps.

Cet accent sur la récurrence du construit et le recours aux échelles de fréquence présupposent généralement que plus les comportements abusifs sont fréquents, plus les effets ressentis sont dévastateurs. Cette présupposition peut toutefois s'avérer fautive (Sasso et Gonzalez-Morales, 2018). Dans une étude menée dans le milieu de l'éducation, Blase, Blase et Du (2008) ont d'ailleurs trouvé que les comportements engendrant le plus grand mal ne sont pas forcément les plus fréquents.

De plus, le concept de récurrence est lui-même mal circonscrit. Certains considèrent que les comportements doivent être fréquents pour être jugés abusifs (Aasland et al., 2010), alors que selon d'autres, ils peuvent survenir occasionnellement (Brunet et al., 2015). Par ailleurs, un superviseur peut avoir recours à une diversité de comportements; chacun d'entre eux pouvant être peu fréquents, et donc risquant de ne pas être considérés abusifs par notre façon de mesurer la variable. Ainsi, le seuil utilisé pour catégoriser un comportement abusif, et le nombre d'items inclus dans le questionnaire, pourraient

grandement influencer les estimations quant à la prévalence du phénomène (ex : Aasland et al., 2010; Brunet et al., 2015).

Objectif 4 : parmi les articles publiés au cours des cinq dernières années, déterminer comment le concept de récurrence est mesuré et quels seuils sont utilisés pour déterminer la prévalence du phénomène.

Exhaustivité et perception de victimisation. Sasso et Gonzalez-Morales (2018), qui étudient le *Workplace mistreatment* et qui abondent dans le même sens que Cortina, Rabelo et Holland (2018), identifient plusieurs lacunes aux outils existants. Ils soulèvent notamment le fait que les comportements contenus dans les instruments ne seraient pas exhaustifs. Le même raisonnement pourrait s'appliquer aux outils mesurant le leadership tyrannique. En ne s'appuyant que sur des mesures comportementales génériques, on risque d'ignorer certaines manifestations abusives (ex : faire du favoritisme ou imposer des demandes déraisonnables).

En outre, Sasso et Gonzalez-Morales (2018) mentionnent qu'en utilisant les mesures actuelles, on considère implicitement les répondants qui subissent les comportements énumérés dans les items des échelles de mesure comme des victimes. Cependant, les répondants ne se considèrent peut-être pas eux-mêmes comme des victimes. En ce sens, deux personnes peuvent indiquer que leur superviseur s'adonne à un certain comportement, une d'elle jugeant celui-ci insignifiant, voire inoffensif, alors que l'autre pourrait le considérer injuste et inacceptable.

Dans ce même ordre d'idées, Sasso et Gonzalez-Morales (2018) soulèvent le fait que les mesures actuelles sont habituellement utilisées de manière à calculer un score moyen global, traitant ainsi chacun des items comme des comportements équivalents les uns aux autres. Ceci peut être problématique, dans la mesure où chaque comportement peut engendrer une intensité de souffrance différente (traduction libre d'*Intensity of harm* de Blase et al., 2008), et aussi dans l'optique où certaines cultures organisationnelles (ex : organisation militaire vs. établissement scolaire) ou ethniques (ex : culture nord-américaine vs. asiatique) pourraient différer sur ce qu'ils considèrent approprié (Tepper et al., 2017; Vogel et al., 2016; Zhang et Liu, 2018). Par ailleurs, plusieurs variables contextuelles pourraient modérer cette perception de victimisation, dont l'intensité du comportement, l'intentionnalité perçue et le pouvoir du superviseur (Herscovis, 2011).

Objectif 5 : identifier les articles publiés au cours des cinq dernières années qui traitent de l'exhaustivité des outils de mesure ou de la perception de victimisation des répondants, ou qui adaptent l'outil utilisé afin de tenir compte de la perspective phénoménologique (Tepper et al., 2017) des répondants.

Ceci rejoint d'ailleurs une citation de Hunter et Schmidt (2004) : « Aucune mesure n'a une validité de construit parfaite; toutes les échelles, même les bonnes, ont au moins une lacune dans leur façon de mesurer le construit (quelque chose d'omis) et une certaine contamination (quelque chose d'inclus qui ne devrait pas l'être) » (traduction libre, p. 1031).

Comparaisons organisationnelles et ethniques. Plusieurs auteurs mentionnent le besoin d'étendre l'étude du réseau nomologique du leadership tyrannique à travers divers milieux organisationnels et ethniques (ex. : Choi et al., 2019; Mackey et al., 2017; Martinko et al., 2013; Tepper, 2007). Au cours des dernières années, plusieurs études abondent en ce sens : la supervision abusive est maintenant étudiée dans des dizaines de pays (Mackey et al., 2017; Martinko et al., 2013) et certaines variables culturelles sont incluses dans les modèles testés empiriquement, telles que l'orientation à la distance au pouvoir, les valeurs traditionnelles et le collectivisme (ex : Bregenzer et al., 2019; Hon et Lu, 2016). Toutefois, peu d'études comparent la manière dont se comportent les instruments de mesures au sein de diverses cultures organisationnelles et ethniques; peu étudient en quoi le réseau nomologique est différent d'une culture à l'autre; et encore trop peu vérifient comment les variables culturelles peuvent avoir un impact sur les perceptions d'abus (Mackey et al., 2017; Martinko et al., 2013).

Objectif 6 : reproduire les analyses effectuées par Mackey et al. (2017), en utilisant les articles publiés entre 2015 et 2020, afin d'examiner à quel point l'étude de la supervision abusive se fait à travers le monde et en quoi les résultats obtenus varient d'un bout à l'autre du globe (en comparant les données des cinq dernières années à celles obtenues par Mackey et al. en 2017); et, repérer toute étude qui compare en quoi un outil génère des résultats différents au sein de diverses cultures organisationnelles ou ethniques, en espérant que certains constats puissent être informatifs quant à l'étude du leadership tyrannique au sein du contexte québécois.

Par ailleurs, même si des échelles mesurant la supervision abusive existent, il s'impose de vérifier si elles ont été validées en contexte éducationnel avant qu'elles soient utilisées dans ce milieu. Ceci vient rejoindre une des recommandations émises par Mackey et ses collègues dans leur méta-analyse de 2017 indiquant un besoin de vérifier comment la dimensionnalité de l'outil se compare d'un type d'emploi à l'autre, d'un type d'organisation à l'autre, ainsi que d'un pays à l'autre; le milieu de l'éducation n'en faisant pas exception.

Objectif 7 : repérer toute étude effectuée sur le leadership tyrannique et la supervision abusive réalisée dans le milieu de l'éducation.

Le secteur de l'éducation

Le système scolaire possède un contexte distinct d'autres types de milieux organisationnels. Il s'agit d'une bureaucratie professionnelle (Mintzberg, 1979) hautement réglementée par les politiques gouvernementales, tout en ayant un centre opérationnel (des professionnels enseignants) qualifié qui exerce un grand contrôle sur l'exécution de son travail et dont le niveau d'interdépendance d'un individu à l'autre est très faible (DeLavergne, 2013, p. 12, repéré à <https://www.univ-montp3.fr/infocom/wp-content/REC-structure-et-dynamique-des-organisations6.pdf>). Hallinger (2003) illustre cette réalité vécue par les directeurs d'établissement scolaire en expliquant qu'ils « occupent un poste de gestion intermédiaire dans lequel leur autorité est sévèrement limitée [par la] nécessité de répondre aux attentes de ceux et celles qui sont hiérarchiquement au-dessus et au-dessous » (traduction libre, p. 335).

Les caractéristiques propres au milieu scolaire pourraient entraîner des comportements tyranniques légèrement différents de la part des leaders. À titre d'exemple, dans certains milieux de travail, un leader abusif pourrait menacer un employé de le renvoyer, mais ce type de menace n'est pas applicable dans les établissements scolaires syndicalisés. Similairement, isoler un employé de ses pairs pourrait être perturbateur, surtout dans un milieu où les échanges sont nécessaires et fréquents entre collègues, mais serait potentiellement moins perturbateur pour des enseignants qui effectuent leur travail de façon très indépendante les uns des autres. Il est donc important de s'assurer qu'une échelle de mesure du leadership tyrannique soit validée en contexte scolaire avant qu'elle soit utilisée par des chercheurs ou des professionnels.

Par ailleurs, la réalité du système scolaire québécois possède elle aussi des particularités qui lui sont propres. Spécialement au cours des deux dernières décennies, le système d'éducation au Québec a connu d'importants changements qui ont modifié la structure, les lois, le curriculum et l'organisation du travail. En 1997, tentant de réduire le fardeau bureaucratique du système, le gouvernement provincial a procédé à une décentralisation des pouvoirs des commissions scolaires vers les établissements d'enseignement, en modifiant la *Loi sur l'instruction publique* (Poirel et Yvon, 2012). Cette initiative a grandement augmenté la charge de travail et les responsabilités des directions d'écoles. De plus, lors de cette même année (1997), le Ministère de l'éducation a implanté une réforme de son curriculum scolaire (Conseil supérieur de l'éducation, 1994), qui a apporté son lot de controverses et a nécessité un réel ajustement de la part des enseignants et des directions scolaires. À cet égard, Poirel et Yvon (2012) écrivent que « depuis la réforme, les directions d'école ont davantage de mesure de contrôle et leur travail s'est alourdi et complexifié » (p. 100).

En revanche, alors que leur charge de travail s'est intensifiée, les directeurs rapportent un manque de ressources influençant négativement leur capacité d'adaptation à ces changements ainsi que leur implantation efficace (Corriveau, 2004). Le système d'éducation a effectivement connu plusieurs coupures budgétaires se traduisant par une diminution du personnel, des services et des ressources disponibles. Autrement dit, on demande aux directions d'école et aux enseignants « d'en faire plus avec moins ».

Dans ce même ordre d'idées, cette réalité que vivent les directeurs d'établissements scolaires rendrait ce milieu de travail davantage propice à la supervision abusive. Selon Tepper et al. (2017), la surcharge de travail des directeurs, la complexification de leurs responsabilités, les objectifs de travail élevés ou irréalistes, la charge émotionnelle liée au traitement des demandes des enseignants et la frustration causée par le manque de ressources octroyées par le gouvernement provincial sont des facteurs qui, au fil du temps, peuvent favoriser le recours à des comportements abusifs. Les théories demandes-ressources de l'emploi de Bakker et Demerouti (2017), de la conservation des ressources de Hobfoll (1989), de la théorie cognitive du stress de Lazarus et Folkman (1984) et de l'*ego depletion* (Baumeister et al., 1998) s'imbriquent entre elles pour expliquer que plus le stress est élevé (résultant d'une perte de ressources, d'une augmentation des demandes et/ou d'un manque de ressources en comparaison aux demandes), plus les ressources du leader s'épuisent et plus sa charge émotionnelle est élevée, plus ses capacités d'auto-régulation diminuent et plus certains leaders (notamment ceux ayant une plus faible capacité de *self-monitoring*) deviennent susceptibles d'utiliser des comportements abusifs (Hobfoll et al., 2018; Tepper et al., 2017). Tel que mentionné, la réalité des administrateurs scolaires comporte des facteurs qui contribuent à diminuer le ratio ressources/demandes et donc qui favoriseraient la perception de supervision abusive chez le personnel. En ce sens, une étude menée par Green (2014) aux États-Unis indique que 90 % des enseignants rapportent être ou avoir été victimes d'un leadership toxique.

Ainsi, ces informations mettent en lumière un réel besoin d'étudier le leadership tyrannique en milieu scolaire, de manière à : (a) complexifier les modèles intégratifs du leadership tyrannique déjà existants, tout particulièrement en améliorant la compréhension des antécédents et des mécanismes qui promeuvent un tel leadership au sein des organisations; (b) effectuer une vérification empirique des connaissances existantes au sein du milieu de l'éducation; ainsi que (c) contrer les effets néfastes associés au leadership tyrannique au sein des établissements scolaires en adaptant les pratiques organisationnelles en matière de sélection et de développement du personnel. Pour y arriver, il est essentiel d'utiliser des outils de mesure de qualité appropriés au contexte étudié (Hinkin, 1998; Mackey et al., 2017). Afin de répondre à ce besoin, une revue systématique de la littérature permettra de (1) recenser les outils de mesure du leadership tyrannique, (2) effectuer un bilan et une analyse de leur qualité méthodologique en répondant aux sept objectifs susmentionnés, et (3) proposer des recommandations aux chercheurs et aux praticiens

œuvrant au sein du milieu scolaire québécois et qui s'intéressent au leadership tyrannique afin de faciliter les choix quant à leur outil.

Méthode

La recension des études a été effectuée au cours de l'été 2020 dans les bases de données *ERIC*, *PsycINFO*, *Web of Science*, *Health and Psychological Instrument* et *Business Source Premier*. Dans l'ensemble des bases de données, à l'exception de la base *ERIC*, les termes *tyran** OR *destructive* OR *abusive* AND *leader** OR *supervis** OR *manager** ont été utilisés. Étant donné qu'*ERIC* est une base de données spécialisée dans le domaine de l'éducation, les termes *administration* OR *administrator** OR *principal** ont été ajoutés aux termes *leader** OR *supervis** OR *manager**, étant appropriés pour cibler les articles portant sur le leadership en contexte scolaire. Afin de prioriser les outils récemment utilisés par la communauté scientifique, seuls les articles publiés au cours des cinq dernières années (2015-2020) ont été retenus.

Sélection des études

Les critères de sélection des études étaient les suivants :

1. Les publications devaient utiliser une échelle quantitative mesurant le leadership tyrannique ou la supervision abusive, tels que ces construits sont définis dans le *Contexte théorique* ci-dessus.
2. L'échelle devait avoir été utilisée en milieu organisationnel, peu importe le type d'organisation.
3. La mesure du leadership tyrannique pouvait être auto-rapportée par le superviseur ou complétée par les subordonnées ou les collègues. Les comportements tyranniques mesurés devaient survenir entre un superviseur/leader et ses employés (et non entre collègues par exemple, comme ce peut être le cas pour le *Workplace Bullying*). Le superviseur ne pouvait pas être fictif, comme c'est le cas parfois dans les devis expérimentaux (ex. : Shum et al., 2020). Les études effectuées sur des étudiants en contexte de formation ont aussi été exclues (ex. : Xu et al., 2018).
4. L'outil devait mesurer des comportements réels et non une intention d'avoir recours aux comportements abusifs (ex. : comme c'est le cas dans la deuxième étude de Liang et al., 2016).
5. Les outils recensés devaient être publiés dans des articles empiriques révisés par les pairs.
6. Seuls les articles rédigés en français ou en anglais ont été retenus.

La méthode employée pour cette revue systématique de la littérature a permis d'identifier les articles empiriques publiés durant les cinq dernières années, dont le protocole mentionne l'utilisation d'un instrument de mesure du leadership tyrannique. Ce sont ces outils de mesure (utilisés dans les articles qui rencontrent les critères de sélection ci-dessus), dont la qualité méthodologique sera analysée et discutée dans le présent travail. Même si la présente étude est centrée sur le leadership tyrannique, les termes

utilisés pour la recherche bibliographique incluaient le leadership destructeur, étant un construit plus large dont le leadership tyrannique fait partie (Aasland et al., 2010; Cacciatore, 2015; Krasikova et al., 2013; Schyns et Schilling, 2013). L'auteure voulait ainsi repérer toute échelle ou sous-échelle pouvant mesurer le construit. Par exemple, l'échelle mesurant le leadership destructeur utilisée par Aasland et al. (2010) comporte une sous-échelle du leadership tyrannique et a donc été sélectionnée.

Qualité méthodologique des outils

Afin de répondre aux sept objectifs susmentionnés et, ultimement, effectuer une analyse de la qualité méthodologique des outils, les informations suivantes ont été extraites des publications recensées :

1. L'instrument de mesure utilisé
2. Le type d'échelle, le nombre de points d'ancrage et le nombre d'items utilisés (objectif 1)
3. La moyenne, l'écart-type et la fidélité (alpha de Cronbach) de l'échantillon tels que rapportés dans la matrice de corrélation de chaque article répertorié (objectif 1)
4. Le nombre de dimension(s) que comporte l'outil (objectif 2)
5. Si (oui ou non) l'article examine empiriquement (ou rapporte des constats empiriques d'une autre étude sur) en quoi les résultats générés par un outil sont comparables à ceux générés par un autre outil mesurant le même construit, que ce soit des corrélations entre les résultats obtenus aux deux questionnaires ou avec le réseau nomologique de la variable (objectif 2)
6. Qui complète le questionnaire (la source : leader, subordonné, collègue) et si (oui ou non) l'étude compare empiriquement les résultats obtenus par diverses sources (objectif 3)
7. En plus du type d'échelle, si (oui ou non) l'échelle utilisée mesure la supervision abusive survenue dans un intervalle de temps passé (et si oui, l'intervalle de temps est noté) et si (oui ou non) un devis longitudinal a été utilisé (objectif 4)
8. Si (oui ou non) l'article évalue la prévalence du leadership tyrannique, et si oui, quel est le seuil utilisé (objectif 4)
9. Si (oui ou non) l'article traite de l'exhaustivité des outils de mesure, de la perception de victimisation des répondants ou adapte sa méthodologie pour en tenir compte (objectif 5)
10. La langue dans laquelle l'outil a été administré, le pays d'échantillonnage, le type d'organisation où la collecte a été effectuée et si (oui ou non) l'article présente des comparaisons entre plus d'une culture organisationnelle ou ethnique au sein d'un même devis (objectifs 6 et 7)

À la suite de l'extraction des éléments susmentionnés, l'auteure du présent travail a procédé à une vérification aléatoire des informations extraites (environ 50 % d'entre elles) afin de repérer toute erreur effectuée lors de la première extraction. Certaines études ayant des informations manquantes n'ont

pu être retenues pour les analyses quantitatives de l'objectif 1, mais ont tout de même été retenues pour la discussion portant sur la qualité méthodologique des outils.

Résultats

Le recours aux termes et aux bases de données mentionnés dans la section *Méthode* a généré 2239 résultats. Le logiciel *Zotero* a permis de repérer les doublons (603 doublons) ainsi que d'effectuer un premier tri des articles basé sur une lecture du titre et du résumé. Ne répondant pas aux critères de sélection, 516 publications ont alors été exclues. Une analyse plus approfondie du contenu des 1120 publications résiduelles a mené à l'exclusion de 850 articles qui ne répondaient pas aux critères de sélection. Au final, 270 articles ont été retenus parmi lesquels on retrouve un total de huit instruments de mesure. La Figure 1 illustre le processus de sélection des publications. Le Tableau 1 énumère les huit outils dont la qualité méthodologique sera discutée dans le présent essai ainsi que le nombre d'études ayant utilisé chaque instrument. Certains articles présentaient plus d'une étude dans une même publication, chacune d'entre elles ayant un échantillon distinct; ceci explique pourquoi le nombre total d'études excède le nombre d'articles sélectionnés. Certains articles ont utilisé le même échantillon dans plus d'une publication; chaque échantillon n'a été retenu qu'une seule fois.

Figure 1

Processus de sélection des articles

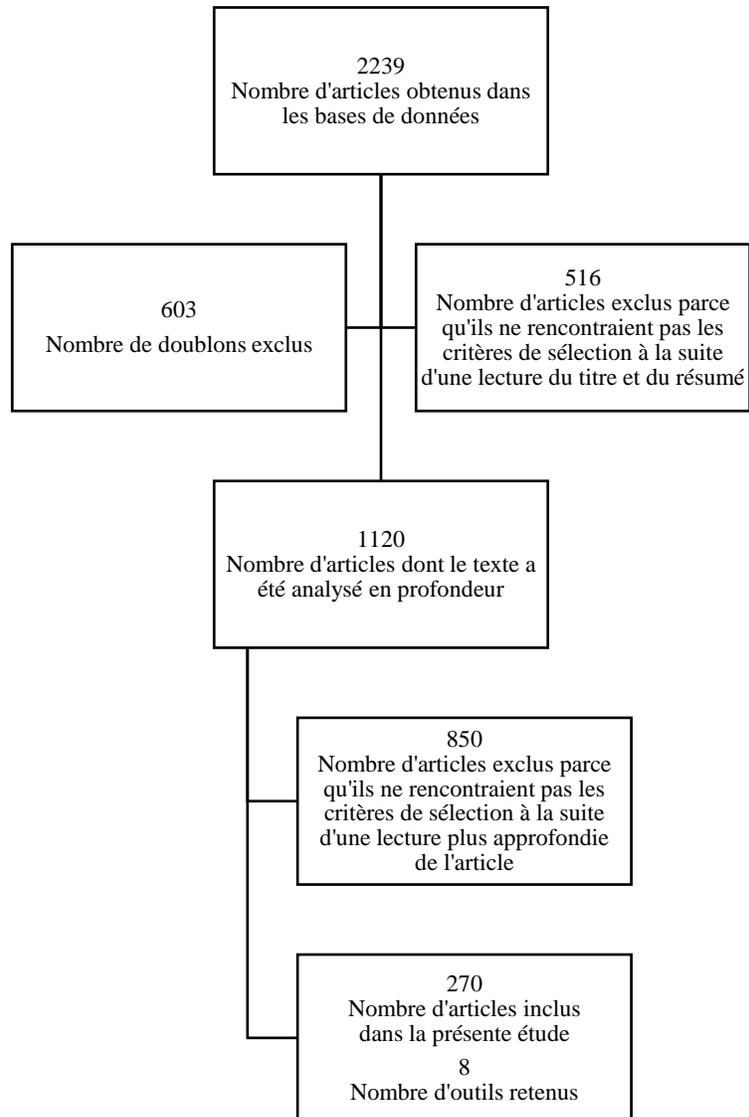


Tableau 1

Liste des échelles incluses dans le cadre du présent essai, le construit qu'elles mesurent et le nombre d'études recensées qui en font l'utilisation

Référence	Construit mesuré	Nombre d'études (<i>k</i>)
1 Tepper (2000) et ses variantes	Supervision abusive	302
Mitchell et Ambrose (2007)	Supervision abusive	85
Aryee et al. (2007)	Supervision abusive	12
Peng et al. (2014)	Supervision abusive	3
Harris, Harvey et Kacmar (2011)	Supervision abusive	1
Lin et al. (2016)	Supervision abusive	1
2 Johnson et al. (2012)	Supervision abusive	5
3 Schmidt (2008)	Leadership toxique	3
4 Aasland et al. (2010)	Leadership destructeur	3
5 Larsson, Brandebo et Nilsson (2012)	Leadership destructeur	3
6 Ashforth (1994)	<i>Petty tyranny</i>	2
7 Cacciatore (2015)	Leadership tyrannique	2
8 Thoroughgood et al. (2012)	Leadership destructeur	1

Note. *k* = 321. Certaines publications présentaient les résultats de plus d'une étude, chacune d'entre elles effectuée sur des échantillons indépendants. Une même étude pouvait aussi utiliser plus d'un outil de mesure du même construit afin de pouvoir effectuer une comparaison des résultats.

Tel que mentionné précédemment, des échelles mesurant le leadership destructeur ou toxique ont été choisies étant donné qu'elles comportent des sous-échelles mesurant le leadership tyrannique. À l'inverse, certains construits sont exclus de la présente étude malgré le fait qu'ils soient très similaires au leadership tyrannique. C'est le cas, par exemple, du leadership exploitant, dont le niveau d'hostilité est moins grand (Schmid, Verdorfer et Peus, 2018). Il est possible d'obtenir des informations quant aux choix qui ont mené à l'exclusion de certains outils recensés dans le cadre de la présente étude en consultant l'Annexe A.

Description des outils recensés

Une compilation d'informations à l'égard des huit outils repérés est disponible au Tableau 2.

Tableau 2

Compilation d'informations extraites à l'égard des outils de mesure recensés

Référence	Qui remplit	Nombre d'items portant sur la SA	Type d'échelle ^a	Nombre d'ancrages	α de l'échelle mesurant la SA	Nombre de dimension(s) attribuées à la SA
1 Tepper (2000) et ses variantes	Subordonnés, collègues et/ou leader	4-15	Fréquence (149) Accord (126)	5-7	0,67-0,98	1-2

2	Johnson et al. (2012)	Subordonnés ou leader	4-5	Fréquence quotidienne (5)	5-6	0,78-0,88	1
3	Schmidt (2008)	Subordonnés	3 ou 7	Accord (3)	5-6	0,76-0,97	1
4	Aasland et al. (2010)	Subordonnés	4	Fréquence (3)	4	0,70-0,76	1
5	Larsson et al. (2012)	Subordonnés	8 ou 12	Fréquence (1)	5-6	0,69-0,93	3
6	Ashforth (1994)	Subordonnés	5	Correct ^b (2)	5	0,81-0,83	---
7	Cacciatore (2015)	Subordonnés	12	Fréquence (2)	3 ou 5	0,93-0,94	3
8	Thoroughgood et al. (2012)	Subordonnés	14	---	5	0,84	1

Note. $k = 321$. ^a Les nombres entre parenthèses représentent le nombre d'articles ayant eu recours au type d'échelle indiqué. ^b Les instructions stipulaient : « *My immediate captain has...* » et, pour chacun des items, les participants répondaient avec une échelle en cinq points, allant de 1 (*not correct*) à 5 (*entirely correct*). SA = supervision abusive; α = coefficient alpha de Cronbach.

Outil de Tepper (2000) et ses variantes. Au cours des cinq dernières années, l'outil de Tepper (2000) est utilisé dans 94 % des études répertoriées (soit un total de 302 études, voir Tableau 1). Parmi celles-ci, seules 54 % d'entre elles (162 études) ont utilisé la version originale à quinze items (telle que décrite dans l'*Introduction*). Parmi les variantes de l'outil de Tepper (2000), c'est la version de Mitchell et Ambrose (2007) qui est la plus populaire (85 études), représentant 28 % des études qui utilisent une version quelconque de l'outil de Tepper (2000). Mitchell et Ambrose (2007) ayant trouvé deux dimensions au construit, leur version de l'instrument est composée de dix des quinze items de l'échelle originale : cinq d'entre eux mesurant les comportements agressifs-actifs, et les cinq autres les comportements agressifs-passifs. La plupart des études (80 des 85 études, 94 % d'entre elles) n'utilisent que cinq items, optant habituellement pour la dimension mesurant les comportements agressifs-actifs. Les répondants évaluent les comportements de leur leader à l'aide d'une échelle d'accord, allant de 1 (fortement en désaccord) à 7 (fortement en accord). La cohérence interne de l'échantillon original (α) était de 0,89.

Aryee et al. (2007) ont sélectionné dix items de l'outil de Tepper (2000) qu'ils trouvaient culturellement neutre; les cinq autres items n'étant pas, selon eux, adéquats au contexte chinois. Les items furent traduits en chinois. Les participants évaluaient les comportements abusifs de leur leader avec une échelle allant de 1 (fortement en désaccord) à 5 (fortement en accord). Ils ont obtenu une cohérence interne (α) de 0,89.

Peng et al. (2014) ont adapté l'outil de Tepper (2000) en prenant trois items de la sous-échelle *active-aggressive* et deux items de la sous-échelle *passive-aggressive* de Mitchell et Ambrose (2007) qu'ils ont traduit en chinois. Les participants notaient la fréquence des comportements abusifs de leur

superviseur envers eux-mêmes en utilisant une échelle allant de 1 (jamais) à 5 (presque toujours). Ils ont rapporté une consistance interne (α) de 0,92. Dans leur devis de recherche, Peng et al. (2014) ont aussi modifié les mêmes cinq items afin de mesurer la supervision abusive dirigée envers les pairs. Pour ce faire, ils ont fait la moyenne des scores de supervision abusive que rapportaient les collègues œuvrant au sein de la même unité de travail (soit chacun des scores que rapportaient les participants concernant les comportements de supervision abusive qu'effectuait leur superviseur envers eux-mêmes, à l'exception du score de l'employé focal). Ils ont obtenu une consistance interne (α) de 0,93 pour la supervision abusive dirigée envers les pairs. Ils ont trouvé que : « le niveau de supervision abusive dont une personne est victime était lié plus négativement avec la performance et l'aide apportée aux collègues lorsque les pairs subissaient des niveaux inférieurs de supervision abusive » (traduction libre, p. 1401).

Lin et al. (2016) ont sélectionné huit items de l'outil de Tepper (2000) susceptibles, selon eux, de varier quotidiennement. Ils ont utilisé l'instrument afin de mesurer la supervision abusive *de base* du leader ainsi que la supervision abusive durant une journée (soit une mesure prise trois jours après la mesure de base) pour tester un modèle de médiation. Pour ce faire, ils ont demandé aux superviseurs d'auto-rapporter la fréquence de leurs comportements abusifs envers leurs employés au cours des quatre semaines précédentes (pour la supervision abusive de base, $\alpha = 0,85$) selon une échelle allant de 1 (jamais) à 5 (chaque jour); ils ont aussi auto-rapportés leur supervision abusive survenue durant la journée ($\alpha = 0,88$) selon une échelle allant de 1 (jamais) à 6 (cinq fois ou plus). Même en ayant utilisé des résultats auto-rapportés par les leaders, Lin et al. (2016) ont obtenu un soutien empirique à leur modèle.

Afin d'ajuster le questionnaire aux besoins des organisations participant à leur étude, Harris, Harvey et Kacmar (2011) ont raccourci la mesure de Tepper (2000) en choisissant six des quinze items. En utilisant les données originales récoltées par Tepper (2000), ils ont obtenu une corrélation de 0,96 en comparant les résultats des deux versions. Les auteurs ont calculé une consistance interne (α) de 0,90 pour le premier échantillon de leur étude et de 0,92 pour le deuxième suggérant, du coup, qu'il serait acceptable d'utiliser un nombre restreint d'items sans influencer la validité de l'outil.

Johnson et al. (2012). Ils ont adapté quatre items de l'échelle de Neumann et Baron (1998) afin de mesurer des comportements agressifs auto-rapportés par les leaders eux-mêmes. Les auteurs ont utilisé l'outil dans le cadre d'un devis longitudinal, lors duquel les mesures de supervision abusives auto-rapportées étaient complétées quotidiennement pendant trois semaines. À la fin de chaque journée, les leaders indiquaient la fréquence de leurs comportements abusifs en utilisant une échelle allant de 1 (jamais) à 6 (cinq fois ou plus). Ce devis longitudinal a entre autres servi à vérifier à quel point les comportements abusifs peuvent varier quotidiennement (la proportion de variance intra-participant rapportée par les auteurs est d'ailleurs chiffrée à 61 %). Même si le construit est traditionnellement mesuré

par les subordonnés (Schyns et Schilling, 2013), Johnson et al. (2012) ont démontré qu'il est possible d'obtenir des liens significatifs entre les variables de leur modèle en sondant les leaders.

Ashforth (1994). Les deux études ayant eu recours à l'outil d'Ashforth dans les derniers cinq ans ont utilisé une version à cinq items (Kant et al. 2013), qui est en fait similaire à celle d'Aasland et al. (2010). Les subordonnés évaluent les comportements de leur leader en utilisant une échelle à cinq ancrages, allant de 1 (pas correct) à 5 (entièrement correct). Les coefficients de fidélité obtenus par les articles recensés sont de 0,81 et 0,83 (Akhtar et Shaukat, 2016; Glaso et al., 2018).

Cacciatore (2015). Cacciatore et ses collaborateurs (2015) ont traduit et validé la mesure d'Ashforth (1994) en français. Cette version de l'outil comporte douze items et présente trois des six dimensions de l'outil original, c'est-à-dire les trois dimensions qui possédaient la meilleure validité discriminante : (1) faire preuve de gestion arbitraire; (2) la dépréciation de ses subordonnés; et (3) appliquer des punitions non contingentes. L'échelle utilisée comporte cinq points d'ancrage (1 = jamais à 5 = fréquemment, sinon toujours). La cohérence interne globale de la version canadienne-française de l'outil est très élevée (alpha de Cronbach de 0,93). En ce qui concerne chacune des dimensions, les alphas de Cronbach sont respectivement de 0,87, 0,91 et 0,90.

Schmidt (2008). L'échelle mesurant le leadership toxique fut développée par Schmidt (2008) au sein d'un milieu militaire (*U.S. Navy* et *U.S. Marine Corps*) en réponse à des lacunes que l'auteur avait observées lors de recherches antérieures. Notamment, des participants indiquaient que, sans être nécessairement la cible de comportements abusifs dirigés envers eux, ils ressentaient une souffrance à être témoin de ceux dirigés envers leurs collègues. La formulation des énoncés en tient compte (ex. : « mon supérieur ridiculise ses subordonnés » plutôt que « me ridiculise », comme c'est le cas dans l'outil de Tepper, 2000). L'auteur argumente aussi que l'échelle de Tepper (2000) fut dérivée de comportements abusifs présents dans des relations romantiques et qu'elle omet certaines formes que peuvent prendre les relations abusives en milieu de travail. D'ailleurs, Schmidt (2008) ajoute quatre dimensions additionnelles à son échelle, en plus de la sous-échelle mesurant la supervision abusive : le *self-promotion*, l'imprévisibilité, le narcissisme et le leadership autoritaire. La sous-échelle de supervision abusive est composée de sept items inspirés de l'outil de Tepper (2000) ainsi que d'incidents critiques identifiés à l'aide de *focus groups*. Les instructions demandent aux participants d'évaluer les comportements de leur superviseur (« *To begin, think of the most destructive supervisor you have experienced and answer each question with regard to this individual.* ») en utilisant une échelle allant de 1 (fortement en désaccord) à 6 (fortement en accord). L'auteur mentionne plusieurs raisons justifiant son choix d'opter pour une échelle d'accord : (1) les échelles de fréquence peuvent être subjectives (ex. : chaque personne peut avoir sa propre conception de ce que « souvent » veut dire; Kline, 2005); (2) les éléments d'ancrage de l'échelle de

fréquence peuvent dépendre du construit mesuré et du contexte (ex. : certains points d'ancrage pourraient ne pas être pertinents; Kline, 2005); et (3) un comportement moins fréquent ne veut pas nécessairement dire qu'il engendre moins d'impact (et donc qu'une échelle de fréquence devrait être accompagnée d'une échelle d'impact; Gatewood, Field et Barrick, 2008). La consistance interne (α) obtenue par Schmidt (2008) pour la sous-échelle de supervision abusive est de 0,93; celle-ci était hautement corrélée ($r = 0,95$, $p < 0,01$) aux résultats générés avec l'échelle originale de Tepper (2000).

Aasland et al. (2010). L'instrument mesurant le leadership tyrannique d'Aasland et al. (2010) s'insère dans un outil mesurant le leadership destructeur basé sur la conceptualisation d'Einarsen et al. (2007), incluant également le *derailed leadership*, le *supportive-disloyal leadership* et le leadership laissez-faire comme construits dits destructeurs. La sous-échelle du leadership tyrannique comporte quatre items auxquels les subordonnés répondent afin de rapporter la fréquence des comportements tyranniques de leur superviseur aux cours des derniers six mois en utilisant une échelle à quatre points d'ancrage, allant de 1 (jamais) à 4 (très souvent/presque toujours). Aasland et al. (2010) ont obtenu une consistance interne (α) pour cette sous-échelle de 0,70.

Larsson et al. (2012). Le *Destrudo-L* fut créé pour mesurer le leadership destructeur en contexte militaire. L'outil est composé de vingt items, regroupés en cinq sous-échelles (soit quatre items par sous-échelle), dont trois chevauchent conceptuellement le leadership tyrannique et la supervision abusive. Selon les auteurs, *Arrogant*, *Unfair* et *Threats*, *Punishments*, *Overdemands* chevaucheraient les deux construits, alors qu'*Ego-oriented*, *False* chevaucherait uniquement la supervision abusive et non le leadership tyrannique. Les subordonnés mesurent les comportements de leur supérieur en utilisant une échelle d'accord à six points. La consistance interne (α) des sous-échelles varient entre 0,80 et 0,84.

Thoroughgood et al. (2012). L'outil mesure les comportements destructeurs des leaders et comporte trois sous-échelles : les comportements dirigés envers les subordonnés (quatorze items), ceux dirigés envers l'organisation (onze items), ainsi que ceux à caractère sexuel (trois items). Les comportements dirigés envers les subordonnés sont ceux qui chevauchent le construit du leadership tyrannique (Einarsen et al., 2007; Krasikova et al., 2013). Dans leur article, Thoroughgood et al. (2012) rapportent une corrélation de 0,69 ($p < 0,01$) entre cette sous-échelle et celle de Tepper (2000). Les subordonnés répondent en utilisant une échelle de fréquence allant de 1 (jamais) à 5 (très souvent). La cohérence interne de la sous-échelle est de 0,96.

Résultats obtenus aux objectifs de recherche

Multiplicité des versions de l'outil de Tepper (2000)

Objectif 1 : L'objectif 1 était de reproduire les analyses portant sur les artéfacts de mesure effectuées par Mackey et al. (2017) en utilisant les articles publiés entre 2015 et 2020, afin de vérifier si le type d'échelle et le nombre d'items utilisés de l'outil de Tepper (2000) sont corrélés à la moyenne et à l'écart-type (obtenus pour la supervision abusive) ainsi qu'à l'indice de fidélité de l'outil rapportés dans les écrits.

Parmi les articles recensés correspondant aux critères de sélection, certains furent exclus pour les analyses permettant de répondre à l'objectif 1, soient ceux ayant des données manquantes ou erronées ainsi que ceux ayant des caractéristiques qui auraient pu ajouter du « bruit », c'est-à-dire des facteurs susceptibles d'expliquer une certaine portion de la variabilité des moyennes (60 échantillons sur les 302 recensés furent ainsi retirés). Par exemple, les échantillons composés d'une population très spécifique et non représentative de la société générale furent exclus ($k = 12$), tout comme les études ayant utilisé une échelle de fréquence dont les libellés des ancrages sont moins communs ($k = 3$, ex. : dont les points d'ancrage sont temporellement définis, comme : *mon superviseur fait le comportement* « une fois par année », « une fois par mois », « une fois par semaine », « une fois par jour », par opposition aux échelles allant de « jamais » à « souvent »). Les données rapportées par une autre source que les subordonnés ($k = 9$) et les études dont la moyenne de supervision abusive représente une mesure quotidienne ($k = 2$) ont aussi été exclues. Ainsi, 242 échantillons indépendants furent utilisés.

Parmi ceux-ci, 56 % ($k = 135$) ont utilisé la version originale de Tepper (2000) à quinze items (ce chiffre s'élevait à 59 % pour les articles publiés avant 2015 recensés par Mackey et al., 2017). Les autres ont opté pour un nombre réduit d'items allant de quatre à quatorze. En moyenne, la communauté scientifique utilise onze items de l'outil original ($\mu = 11,17$, $\sigma = 4,62$), ceci est le même constat que celui obtenu par Mackey et al. en 2017. En ce qui concerne le type d'échelle, 49 % ($k = 119$) ont utilisé une échelle de fréquence, 43 % ($k = 103$) une échelle d'accord, et 8 % ($k = 20$) ne l'ont pas spécifié. En comparaison aux articles publiés avant 2015, le recours aux échelles d'accord semble avoir gagné en popularité; Mackey et al. (2017) avaient identifié deux fois plus d'études qui utilisaient une échelle de fréquence ($k = 84$) en comparaison à celles ayant recours à une échelle d'accord ($k = 46$). Comme c'était le cas pour les articles publiés avant 2015, le nombre de points d'ancrage est variable : 67 % ($k = 162$) ont opté pour cinq, 5 % ($k = 13$) pour six, 23 % ($k = 55$) pour sept et moins de 1 % ($k = 1$) pour neuf.

Les coefficients de corrélations obtenus dans le cadre du présent travail (présentés au Tableau 3) sont aussi très comparables à ceux obtenus en 2017 (Tableau 4). Le *nombre d'items* utilisés de l'outil de

Tepper (2000) est peu corrélé à la moyenne ($r = 0,01$, $p = 0,87$) et à l'écart-type ($-0,11$, $p = 0,10$), mais modérément corrélé à la cohérence interne alpha de Cronbach ($r = 0,32$, $p < 0,001$). En revanche, peu importe le nombre d'items (et même le type d'échelle) utilisé, les indices de fidélité rapportés dans les écrits sont habituellement élevés et ne posent pas de problème ($\mu = 0,91$, $\sigma = 0,06$). Concernant le *type d'échelle* (codé par 0 pour les échelles de fréquence et par 1 pour les échelles d'accord, comme l'avaient fait Mackey et al. en 2017), celui-ci est moyennement corrélé à la moyenne obtenue dans les écrits ($r = 0,30$, $p < 0,001$), mais peu corrélé à l'écart-type ($r = 0,11$, $p = 0,10$) et à la cohérence interne ($r = -0,04$, $p = 0,57$). La direction de la corrélation indique que les moyennes générées par l'utilisation d'une échelle d'accord seraient plus élevées que celles obtenues à l'aide d'une échelle de fréquence.

Tableau 3

Artéfacts de mesure de la perception de supervision abusive du présent essai

Variable	1	2	3	4	5	μ	σ
1. Moyenne mesurée	----					1,95	0,70
2. Écart-type mesuré	0,39***	----				0,79	0,45
3. Nombre d'items	0,01	-0,11	----			11,17	4,62
4. Type d'échelle	0,30***	0,11	-0,16*	----		0,46	0,50
5. Consistance interne	-0,21***	0,17**	0,32***	-0,04	----	0,91	0,06

Note. $k = 222-242$. Le type d'échelle a été codé par 0 pour les échelles de fréquence et par 1 pour les échelles d'accord. * $p < 0,05$ ** $p < 0,01$ *** $p < 0,001$ μ = moyenne agrégée des données des échantillons; σ = écart-type de la moyenne agrégée des données des échantillons.

Tableau 4

Artéfacts de mesure de la perception de supervision abusive obtenus par Mackey et al. (2017)

Variable	1	2	3	4	5	μ	σ
1. Moyenne mesurée	----					1,78	0,46
2. Écart-type mesuré	0,72	----				0,81	0,30
3. Nombre d'items	-0,03	-0,07	----			11,83	4,37
4. Type d'échelle	0,29	0,28	-0,45	----		0,35	0,48
5. Consistance interne	0,19	0,40	0,29	-0,09	----	0,92	0,04

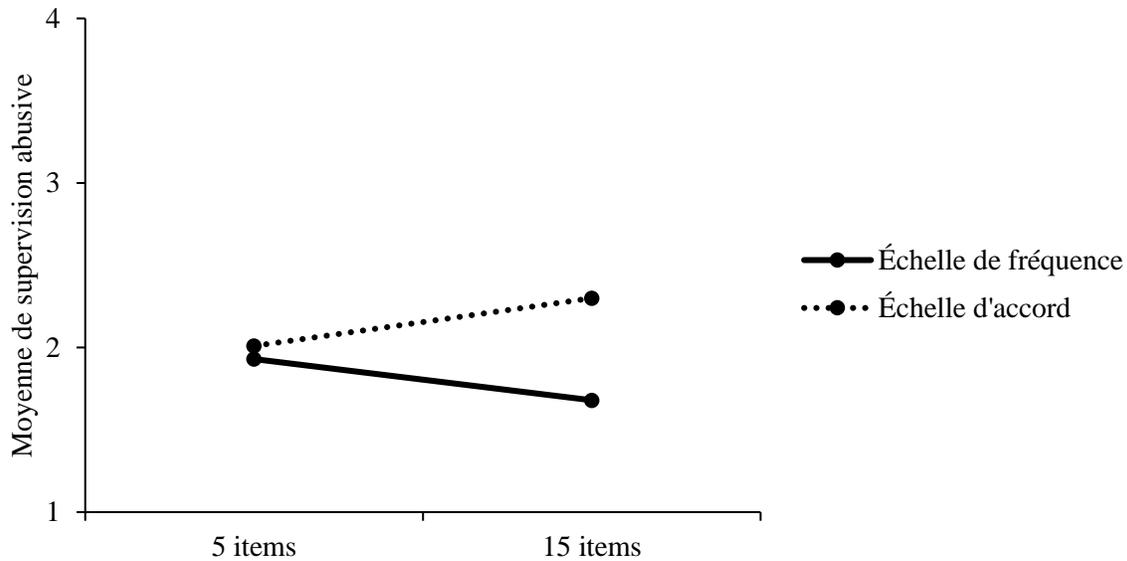
Note. $k = 130$. Le type d'échelle a été codé par 0 pour les échelles de fréquence et par 1 pour les échelles d'accord. μ = moyenne agrégée des données des échantillons; σ = écart-type de la moyenne agrégée des données des échantillons.

Cela dit, en faisant les analyses vouées à vérifier les postulats requis pour effectuer des analyses corrélationnelles (Yergeau et Poirier, 2013), on peut constater que le postulat de l'homoscédasticité des résiduels n'est pas rencontré. En effet, les participants optent majoritairement soit pour la version de l'outil à cinq items, soit pour celle à quinze items; la variabilité des résiduels est donc inégale. Par ailleurs, dans le but de reproduire la méthodologie de Mackey et al. (2017), le type d'échelle avait été codé en variable continue, mais il s'agit réellement d'une variable catégorique (les chercheurs utilisant soit une échelle d'accord ou de fréquence). Ainsi, une analyse de variance (ANOVA avec devis factoriel; Abbott, 2016) fut effectuée, permettant de déterminer si (a) l'effet principal de chacune des variables indépendantes (le nombre d'items et le type d'échelle) sur la variable dépendante (moyenne de supervision abusive rapportée dans les écrits) est significatif, et si (b) l'interaction entre les variables indépendantes (le nombre d'items et le type d'échelle) influence significativement la variable dépendante (la moyenne de supervision abusive). Une telle ANOVA permet de vérifier si la différence observée entre les résultats obtenus pour divers groupes est réellement significative ou due à l'erreur d'échantillonnage, étant donné les différences entre les études.

L'ANOVA indique que le modèle généré est significatif ($\alpha = 0,05$, $F(3, 193) = 8,53$, $p < 0,001$). L'effet d'interaction entre le nombre d'items et le type d'échelle utilisés est significatif ($F(1, 193) = 7,19$, $p < 0,01$); sa taille d'effet est relativement petite ($\eta^2 = 0,04$; Cohen, 1988). Ce qui est intéressant, c'est que l'effet simple du type d'échelle n'est pas significatif lorsqu'on a recours à la version à cinq items ($F(1, 193) = 0,23$, $p = 0,63$, $\eta^2 = 0,00$), mais elle l'est lorsqu'on utilise la version à quinze items ($F(1, 193) = 25,19$, $p < 0,001$, $\eta^2 = 0,12$); sa taille d'effet étant relativement grande. Cette différence est illustrée à la Figure 2. Les moyennes et les tailles d'échantillons sont présentées à l'Annexe B. Dans cette même annexe, on peut également y retrouver des analyses additionnelles montrant notamment que les vingt-cinq études ($k = 25$) exclues, ayant eu recours à un autre nombre d'items que cinq ou quinze, ne diffèrent pas des études retenues pour l'ANOVA (lorsqu'on examine les caractéristiques recensées dans le présent travail). On peut également y trouver une discussion sur les postulats et la puissance statistique des analyses.

Figure 2

Moyenne de supervision abusive selon le nombre d'items de l'outil de Tepper (2000) et le type d'échelle utilisés



Ainsi, si on ne se fie qu'aux coefficients de corrélation, les constats effectués par Mackey et al. en 2017 demeurent les mêmes en analysant les articles publiés au cours des cinq dernières années. Premièrement, l'utilisation d'un nombre réduit d'items ne semblerait pas problématique en soi. Deuxièmement, le type d'échelle pourrait influencer les résultats obtenus lorsque nous avons recours à l'outil de Tepper (2000), de sorte que les moyennes générées par une échelle d'accord auraient tendance à être plus élevées que celles générées par une échelle de fréquence. Néanmoins, l'ANOVA effectuée apporte une précision, comme quoi le choix du type d'échelle importerait peu lorsqu'on utilise la version courte de l'outil de Tepper (2000) à cinq items, mais la différence observée (moyennes plus élevées lorsqu'on utilise une échelle d'accord versus de fréquence) est significative lorsqu'on utilise la version originale à quinze items. Enfin, la moyenne globale de la supervision abusive obtenue pour l'ensemble des échantillons est relativement faible ($\mu = 1,95$, $\sigma = 0,70$), supportant ainsi le fait que la supervision abusive soit considérée comme un phénomène peu fréquent (Mackey et al., 2017).

Multiplicité des conceptualisations

Objectif 2 : L'objectif 2 consistait à identifier quelles sont les diverses façons de concevoir la dimensionnalité du construit et à repérer les articles qui examinent empiriquement en quoi cette diversité de façons de concevoir le leadership tyrannique permet de mieux prédire (ou de prédire différemment) le réseau nomologique de la variable.

Même si la vaste majorité des études quantitatives effectuées au cours des cinq dernières années utilisent l'échelle de Tepper (2000) ou une variante de celle-ci, 6% des études ($k = 19$) ont recours à d'autres outils pour mesurer le leadership tyrannique (ex. : Aasland et al., 2010; Boudrias et al., 2020; Schmidt, 2008). Certaines des échelles recensées, décrites ci-dessus, présentent la supervision abusive comme une dimension s'inscrivant dans un construit de leadership négatif plus large, comme le leadership destructeur ou le leadership toxique (Aasland et al., 2010; Larsson et al., 2012; Schmidt, 2008; Thoroughgood et al., 2012). Les écrits utilisent habituellement un outil qui conçoit le construit en une seule dimension (Johnson et al., 2012; Tepper, 2000), mais plusieurs (27 % des écrits recensés, $k = 85$) ont recours à l'échelle de Mitchell et Ambrose (2007) qui en identifie deux : les comportements abusifs passifs-agressifs et actifs-agressifs. La présente recension a également permis de repérer une étude (Fatima et Bashir, 2017) qui a obtenu un soutien empirique à la double-dimensionnalité décrite par Tepper, Duffy et Breaux-Soignet (2012), soit : le type impulsif (actions automatiques et non contrôlées) et le type stratégique (actions délibérées ayant un but particulier). D'autres échelles attribuent également plusieurs dimensions au leadership tyrannique, dont celle d'Ashforth (1994) qui en présente six, ainsi que celle de Cacciatore (2015) et de Boudrias et al. (2020) qui en présente trois. Le nombre de dimensions que comporte chaque échelle figure dans le Tableau 2.

Toutefois, parmi les articles recensés, très peu comparent empiriquement en quoi les données obtenues par les divers outils sont similaires et/ou distinctes. Liao et al. (2020) ont trouvé une corrélation de 0,78 entre les données de supervision abusive mesurées « au quotidien » par l'outil de Johnson et al. (2012) et les données de supervision abusive chronique générées par l'outil de Tepper (2000). Similairement, Yu et Duffy (2020) ont calculé une corrélation de 0,69 ($p < 0,01$) entre les données obtenues à l'aide de l'outil de Johnson et al. (2012; comportements abusifs mesurés « au quotidien ») et celles générées par l'instrument de Tepper (2000; comportements abusifs survenus « au cours des derniers deux mois »). Schmid et al. (2019) ont obtenu une corrélation significative frôlant la singularité ($r = 0,86$, $p < 0,001$) entre l'échelle de Tepper (2000) et la sous-échelle des comportements destructeurs dirigés envers les subordonnés de Thoroughgood et al. (2012). Tel que mentionné précédemment, Boudrias et al. (2020) rapportent que l'outil de Cacciatore (2015) à trois dimensions est hautement corrélé ($r = 0,77$) à celui de Tepper (2000) à une seule dimension. Aucune des études recensées ne comparent cependant empiriquement la validité prédictive ou incrémentielle d'un outil par rapport à un autre.

Deux études ajoutent un bémol non négligeable à l'égard de la dimensionnalité du construit. D'une part, Breevhart et de Vries (2017) ont testé un modèle qui considérait la supervision abusive comme une variable dépendante, et les domaines de la personnalité (HEXACO-PI-R) comme des variables indépendantes. De façon intéressante, leur modèle expliquait près du double de la variance de la

supervision abusive lorsque celle-ci était mesurée par la version à cinq items ($R^2 = 0,65$, $p < 0,001$) de Mitchell et Ambrose (2007, comportements agressifs-actifs) en comparaison à la version à quinze items ($R^2 = 0,35$, $p < 0,001$) de Tepper (2000). Si la supervision abusive s'avère être un construit à une seule dimension, les auteurs proposent qu'il serait peut-être mieux capté avec une mesure à cinq items qu'à quinze items. D'ailleurs, Breevhart et de Vries (2017) n'ont pu reproduire les résultats de Mitchell et Ambrose (2007) en ce qui concerne la double-dimensionnalité du construit. D'autre part, Pundt et Schwarzbeck (2018) ont testé un modèle qui conceptualisait la supervision abusive comme la variable dépendante et trois autres variables (le *supervisor irritation*, le *supervisor self-control* et l'*external monitoring*) comme les variables indépendantes. Contrairement à ce qu'ont obtenu Breevhart et de Vries (2017), la variance expliquée de leur modèle était plus grande lorsque la supervision abusive était mesurée avec les treize items ($R^2 = 0,51$, $p < 0,01$) que lorsqu'elle était mesurée par la sous-échelle des comportements agressifs-actifs ($R^2 = 0,36$, $p < 0,01$) ou par celle des comportements passifs-agressifs ($R^2 = 0,37$, $p < 0,01$). Les constats de Breevhart et de Vries (2017) et de Pundt et Schwarzbeck (2018) apportent une nuance importante, comme quoi la validité du construit pourrait être affectée par le nombre d'items utilisés; les deux études rapportent néanmoins des constats inverses.

Dans l'étude de Pundt et Schwarzbeck (2018), les résultats obtenus avec l'outil de Tepper (2000; en utilisant treize des quinze items) étaient hautement corrélés à ceux obtenus avec les deux sous-échelles à cinq items de Mitchell et Ambrose publiée en 2007 ($r = 0,78$, $p < 0,01$ pour les comportements agressifs-actifs; $r = 0,77$, $p < 0,01$ pour les comportements agressifs-passifs). La corrélation entre les deux sous-échelle de Mitchell et Ambrose (2007; $r = 0,52$, $p < 0,01$) était de moins grande envergure, comme quoi les résultats des deux sous-échelles seraient davantage corrélées avec l'outil de Tepper (2000) qu'entre elles. Ceci semble indiquer que les deux sous-échelles mesureraient des facettes différentes du construit. Il se pourrait que les relations nomologiques puissent être différentes pour les deux dimensions. Les auteurs notent aussi que la cohérence interne (α) de la sous-échelle des comportements passifs-agressifs n'est habituellement pas rapportée dans les écrits; celle-ci serait faible (elle était de 0,58 pour l'échantillon de leur étude). Ceci explique possiblement pourquoi la majorité des études répertoriées dans le présent essai, qui ont recours à la variante de Mitchell et Ambrose (2007), optent pour n'utiliser que la dimension des comportements agressifs-actifs. Pundt et Schwarzbeck (2018) en concluent que la communauté scientifique sous-estimerait probablement la relation entre certaines variables et la forme passive de la supervision abusive, compte tenu de la faible fiabilité de la mesure associée à cette dimension.

Néanmoins, ces quelques résultats recensés montrent que très peu d'études examinent en quoi les diverses façons de concevoir le leadership tyrannique sont liées entre elles et en quoi cette diversité permet de mieux prédire (ou de prédire différemment) le réseau nomologique de la variable.

Diversité de sources

Objectif 3 : L'objectif 3 visait à identifier, dans les écrits publiés entre 2015 et 2020, par qui (leaders, subordonnés et/ou collègues) les comportements de supervision abusive sont habituellement mesurés, ainsi qu'à repérer les études qui comparent empiriquement en quoi les données de supervision abusive générées par les outils varient selon la source à qui on les administre.

La plupart des outils sont strictement complétés par les subordonnés qui évaluent les comportements de leur leader (Aasland et al., 2010; Ashforth, 1994; Cacciatore, 2015; Larsson et al., 2012; Schmidt, 2008; Thoroughgood et al., 2012). Deux études ayant recours à l'outil de Johnson et al. (2012) et sept études ayant recours à celui de Tepper (2000; ou à une variante) ont demandé aux leaders de mesurer leurs comportements abusifs de façon auto-rapportée (mesures quotidiennes, devis longitudinal). Les auteurs ayant fait ce choix (ex. : Courtright et al. 2016; Johnson et al., 2012; Yeung et Shen, 2019) expliquent que les comportements des leaders peuvent être différents d'un employé à l'autre. Les comportements mesurés par un subordonné peuvent donc ne pas refléter à quel point un leader peut avoir été abusif avec d'autres subordonnés au cours d'une journée. Même si les données auto-rapportées par les leaders risquent d'être plus élevées que celles rapportées par les employés, les auteurs argumentent que les leaders donnent ainsi une mesure plus réelle de leurs comportements abusifs (Yeung et Shen, 2019). Johnson et al. (2012) mentionnent aussi que les données fournies par les subordonnés peuvent être biaisées par leurs théories implicites du leadership (Epitropaki et Martin, 2004) et par leurs perceptions quant à la prototypicalité du groupe (van Knippenberg, 2011). Il serait donc approprié d'avoir recours aux données auto-rapportées par les leaders, d'autant plus si les autres variables incluses dans l'étude sont, elles aussi, centrées sur le leader (*leader-centric approach; within-person relationships*; Liao et al., 2018; Yeung et Shen, 2019).

D'ailleurs, Berry et al. (2012) ont réalisé une méta-analyse portant sur les comportements contreproductifs indiquant que les données mesurées par diverses sources (*self- and other-ratings*) sont grandement corrélées entre elles et qu'elles engendreraient un pattern corrélationnel similaire avec des variables du réseau nomologique. Ceci dit, selon les articles recensés par la présente étude, une telle méta-analyse n'aurait pas été faite pour les données de supervision abusive.

Supervision abusive vicariante. Dans quatorze études, l'instrument de Tepper (2000 ou une de ses variantes) a aussi été adapté afin de mesurer la supervision abusive vicariante, c'est-à-dire les comportements abusifs manifestés par le leader envers un ou des collègue(s) autre que soi (ex. : Huang et al., 2019; Mitchell et al., 2015; Peng et al., 2019; Waldman et al., 2018; Wee et al., 2017). Plusieurs auteurs ont obtenu une corrélation élevée entre la supervision abusive « envers soi » et « envers les pairs »

($r = 50$, $p < 0,01$, Chen et Liu, 2019; $r = 0,60$ $p < 0,001$, Jiang et al., 2019; $r = 0,54$, $p < 0,01$, Kluemper et al., 2019; $r = 0,68$ et $0,38$, $p < 0,01$, Priesemuth et Schminke, 2019 dans leurs deux études; $r = 0,65$ et $0,81$, $p < 0,01$, Shao et al., 2018 dans deux études différentes). Parmi ces quelques études, les moyennes de la supervision abusive dirigée « envers les pairs » tendent à être plus élevées que celles des comportements dirigés « envers soi » (Chen et Liu, 2019; Priesemuth et Schminke, 2019; Shao et al., 2018), tout en gardant à l'esprit qu'aucun test d'hypothèse ne fut effectué à cet égard.

Au Canada, Ogunfowora, Weinhart et Hwang (2019) ont obtenu une corrélation de $0,71$ ($p < 0,01$) entre la supervision abusive effectuée « envers soi » et celle « envers les pairs » (en ayant recours à sept items de l'outil de Tepper, 2000). Dans cette même étude, les auteurs montrent aussi un soutien empirique à la validité incrémentielle de l'*abusive supervision differentiation* (c'est-à-dire à quel point le superviseur manifeste des comportements abusifs envers certains subordonnés et pas d'autres) en comparaison à la supervision abusive dirigée « envers soi ».

Zhang, Liu et Chen (2019) ont également adapté les outils de Tepper (2000) et de Mitchell et Ambrose (2007) afin de mesurer la supervision abusive vicariante. En suivant la même procédure que Kiewitz et al. (2012) pour valider la version raccourcie d'un questionnaire, ils ont comparé les résultats de supervision abusive vicariante obtenus à l'aide des versions à cinq et à quinze items. Comme les résultats des deux versions étaient hautement corrélés ($r = 0,96$, $p < 0,01$), les auteurs en concluent que la version courte de l'échelle est valide pour mesurer la supervision abusive manifestée « envers les pairs ».

Sans que leur étude porte sur la supervision abusive vicariante, Lukacik et Bourdage (2019) ont voulu valider les données de supervision abusive fournies par les employés en administrant l'outil de Tepper (2000) à un collègue de travail (et ce pour chacun des répondants) afin qu'il/elle mesure les comportements du superviseur envers l'employé en question. Les données provenant des deux sources étaient très corrélées ($r = 0,79$).

Quelques études ont modifié l'outil de Tepper (2000) afin de mesurer les comportements dirigés envers l'équipe (Jiang et Gu, 2016; Rousseau et Aubé, 2018; Samreen, et Rashid, 2018; Schaubroeck, Peng et Hannah, 2016; Yam et al., 2016), permettant l'étude du phénomène au niveau groupal.

Les constats susmentionnés montrent que peu d'études explorent empiriquement à quel point les résultats de supervision abusive auto-rapportés par les leaders sont réellement comparables à ceux mesurés par les employés. Les données générées par les outils ciblant les comportements manifestés « envers soi » semblent grandement corrélées à celles obtenues par les outils mesurant les comportements dirigés « envers les collègues ». Les tests d'hypothèses appuyant la direction des comparaisons demeurent toutefois rares et les facteurs influençant ces corrélations restent peu étudiés.

Réurrence et prévalence

Objectif 4 : parmi les articles publiés dans les cinq dernières années, l'objectif 4 visait à déterminer comment le concept de réurrence est mesuré et quels seuils sont utilisés pour déterminer la prévalence du phénomène.

Réurrence et devis transversaux. La réurrence des comportements abusifs est mesurée de diverses façons tout dépendant de l'outil et du type de devis utilisés. Parmi les publications répertoriées, 92 % des études (k = 292) mesurent la supervision abusive à un point dans le temps en demandant, par exemple, de noter les comportements du superviseur dans un intervalle de temps passé. Parmi ces études, le concept de réurrence est mesuré en utilisant une échelle de fréquence (outils d'Aasland et al., 2010, Cacciatore, 2015, Larsson et al., 2012, Tepper, 2000 et ses variantes), et parfois, en utilisant une échelle d'accord (outils de Schmidt, 2008, Tepper, 2000 et ses variantes). Dans trois études (Kernan, Racicot et Fisher, 2016; Low, Sambasivan et Ho, 2019; Richard et al., 2020), les auteurs ont opté pour une échelle de fréquence dont les ancrages sont temporellement définis, allant respectivement : de 0 (jamais) à 4 (quotidiennement) dans l'étude de Kernan et al. (2016); de 1 (jamais) à 5 (hebdomadairement) dans celle de Low et al. (2019); et de 1 (jamais) à 7 (quotidiennement) dans celle de Richard et al. (2020). L'utilisation de tels ancrages temporellement définis permettraient de réduire l'ambiguïté des échelles de fréquence habituellement utilisées (ex. : quelle est la différence entre « souvent » c. « très souvent » de Tepper, 2000).

Réurrence et devis longitudinaux. Toutefois, certaines études optent pour un devis longitudinal. Onze études ont administré le questionnaire de Tepper (2000, ou une variante) à plus d'un temps de mesure, chaque collecte étant séparée par plusieurs semaines ou mois; et une étude (Nielsen et al., 2016) a fait l'équivalent en ayant recours au questionnaire d'Aasland et al. (2010). Seize autres études ont opté pour un devis longitudinal « au quotidien », c'est-à-dire des mesures journalières effectuées sur une, deux ou trois semaines de travail (onze ont utilisé l'outil de Tepper, 2000, et cinq, l'outil de Johnson et al., 2012). En comparaison, Martinko et al. ont repéré une seule étude sur soixante-deux en 2013 ayant utilisé un *vrai* devis longitudinal (par opposition aux *multi-waves studies*). Le nombre de devis longitudinaux ayant augmenté, il s'agit d'une amélioration, surtout qu'ils permettent de mieux examiner certaines relations causales (Martinko et al., 2013) ainsi que les variations temporelles des comportements abusifs (Bormann, 2017).

Réurrence et prévalence. Au cours des cinq dernières années, deux études ont estimé la prévalence du phénomène. Une d'entre elles fut réalisée en Norvège au sein de l'industrie pétrolière en ayant recours à la mesure d'Aasland et al. (2010; Nielsen et al., 2016). À l'aide de l'échelle à quatre points

d'ancrage décrite précédemment, un comportement devait survenir *quite often* ou *very often* afin d'être considéré destructeur (*operational criterion method*). En utilisant ce seuil, Nielsen et al. (2016) ont estimé que 4,21 % des répondants ont été exposés à au moins un comportement de leadership tyrannique dans les six mois précédant l'étude. Au Québec, Brunet et al. (2015) ont estimé la prévalence de ce type de leadership au sein d'une population d'infirmières en utilisant l'outil de Cacciatore (2015). À l'aide d'une échelle de fréquence à trois points d'ancrage, 34 % des répondants ont indiqué être occasionnellement victimes d'au moins un comportement de leadership tyrannique, alors que 9 % ont rapporté l'être régulièrement.

Ainsi, le recours aux devis longitudinaux serait à la hausse (en comparaison au nombre de devis longitudinaux recensé par Martinko et al. en 2013) et il s'agirait de la façon idéale permettant de capter la récurrence du phénomène. Cela dit, à défaut de pouvoir toujours opter pour un devis longitudinal, il ne semble pas y avoir de consensus quant au type d'échelle à utiliser dans un devis transversal. L'échelle de fréquence permettrait d'estimer la prévalence du phénomène et de capter une certaine récurrence des comportements (ex. : jamais c. souvent), mais elle peut être sujette à interprétation par les participants. L'échelle d'accord peut être moins sujette à interprétation, mais elle ne comporte pas d'ancrage temporel. Même si les ancrages temporellement définis précisent, en quelque sorte, les ancrages de fréquence, le risque de ne pas capter l'intensité, la souffrance et l'effet cumulatif des comportements qui surviennent moins fréquemment demeure.

Exhaustivité et perception de victimisation

Objectif 5 : l'objectif 5 consistait à identifier les articles publiés au cours des cinq dernières années qui traitent de l'exhaustivité des outils de mesure, de la perception de victimisation des répondants et/ou qui adaptent un instrument de mesure existant en fonction de celles-ci.

Aucune des études recensées (utilisant un outil quantitatif) ne traite empiriquement de l'exhaustivité des instruments. Il faudrait probablement examiner les études qualitatives à cet effet. Quoique la présente recension des écrits portait sur les articles quantitatifs mesurant la supervision abusive, aucun article qualitatif ne fut repéré, lors de l'examen du corpus de la littérature retenu dans cet essai (cf. p. 25), traitant de l'exhaustivité des instruments.

Cependant, quelques articles recensés tiennent compte de la perception de victimisation en ayant recours à la méthode des incidents critiques (Flanagan, 1954). Cette méthode permettrait de recueillir plus d'informations sur la nature des événements abusifs ainsi que sur les perceptions et les réactions des présumées victimes (Herschovich, 2011). Elle consiste à demander aux participants de se remémorer un incident *abusif* et de répondre aux questionnaires de l'étude en ayant l'exemple en tête. Il est possible que

la méthode influence les émotions ressenties par les répondants au moment où ils complètent les questionnaires (Thomas et Diener, 1990), mais le fait de repenser aux détails contextuels de l'événement permettrait une plus grande exactitude et une plus grande vivacité du souvenir (Lang et al., 1980; Robinson et Clore, 2001), ce qui permettrait de mieux capter le phénomène par les instruments de mesure. Mitchell et al. (2015) et Priesemuth et Schminke (2019) ont eu recours à cette méthode pour étudier les effets de la supervision abusive vicariante sur certaines croyances, émotions et comportements des collègues de travail (tierce partie observatrice). Priesemuth et Bigelow (2020) ont également utilisé cette méthode en demandant au leader lui-même de se rappeler un événement où il a manifesté des comportements abusifs envers un employé. Par ailleurs, Kim et al. (2019) ont utilisé la méthode des incidents critiques pour explorer la perception des motifs qu'attribuent les subordonnés aux comportements abusifs de leur leader, en tant que modérateur de la relation entre la supervision abusive et la qualité de la relation employeur-employé.

Cela dit, dans un tel cas, le score de supervision abusive obtenu ne porte habituellement que sur cet événement spécifique. Il serait potentiellement plus difficile de capter l'aspect de récurrence du construit (ceci n'étant pas le but visé par la méthode). Par ailleurs, étant donné qu'elle a été utilisée dans un nombre très restreint d'études, il reste à voir en quoi les scores obtenus par les instruments mesurant la supervision abusive en sont influencés.

En outre, Tepper et al. (2015) conceptualisent la perception de victimisation comme une variable en soi. Dans leur étude, ils la traitent comme une variable médiatrice (entre les comportements hostiles et certaines variables d'intérêt), qu'ils ont nommé *victim identity*, et ont créé un questionnaire pour la mesurer.

Comparaisons organisationnelles et ethniques

Objectif 6 : l'objectif 6 consistait à reproduire les analyses effectuées par Mackey et al. (2017), en utilisant les articles publiés entre 2015 et 2020, afin d'examiner à quel point l'étude de la supervision abusive se fait à travers le monde et en quoi l'outil de Tepper (2000) génère des résultats différents d'un bout à l'autre du globe (en comparaison aux constats que Mackey et al. avaient publiés dans leur analyse des artéfacts de mesure en 2017); ainsi qu'à repérer toute étude qui compare empiriquement les résultats obtenus dans plus d'une culture organisationnelle ou ethnique dans un même devis.

Au cours des cinq dernières années, la supervision abusive a été étudiée dans une trentaine de pays à travers le monde (Tableau 5). Tel que mentionné par Mackey et al. en 2017, les moyennes de supervision abusive (obtenues avec l'outil de Tepper, 2000) rapportées par les pays orientaux ($\mu_{AS} = 2,25$, $\sigma_{AS} = 0,78$, $k = 120$) tendent à être plus élevées que celles publiées par les pays occidentaux ($\mu_{AS} = 1,57$,

$\sigma_{AS} = 0,40$, $k = 81$). En utilisant la même base de données qu'à l'objectif 1, une ANOVA à un facteur montre une différence significative entre les moyennes observées dans ces deux régions du globe ($F(1,199) = 51,90$, $p < 0,001$) et sa taille d'effet est grande ($\omega^2 = 0,20$; Cohen, 1988).

Tableau 5

Résultats obtenus avec l'outil de Tepper (2000) par pays

Pays	N	k	μ_{SA}	σ_{SA}	Nombre d'items	Type d'échelle	r_{xx}
Allemagne	1911	7	1,50	0,64	12,57	0,29	0,91
Allemagne et États-Unis	150	1	1,73	0,84	10,00	----	0,93
Amérique du Nord	302	2	1,33	0,49	10,00	0,00	0,92
Angleterre	1 621	3	1,68	0,93	15,00	0,33	0,95
Autriche, Allemagne et Slovénie	503	1	2,09	1,08	13,00	0,00	0,98
Belgique	225	2	1,42	0,55	10,00	1,00	0,82
Canada	1 600	2	1,53	0,79	5,00	1,00	0,94
Chine	25 918	77	2,11	0,79	9,97	0,54	0,89
Émirats arabes unis	335	1	1,95	0,77	5,00	----	0,83
Équateur	305	1	2,33	1,33	6,00	1,00	0,84
États-Unis	13 426	55	1,53	0,75	11,49	0,20	0,93
Finlande	262	1	1,44	0,89	5,00	1,00	0,91
France	205	1	1,82	0,94	15,00	0,00	0,94
Hollande	411	3	1,71	0,89	8,33	1,00	0,91
Inde	3 402	8	2,96	0,82	12,50	0,67	0,90
Israël		1					
Japan	1 800	3	2,05	0,85	15,00	0,00	0,97
Korée du Sud	2 939	13	1,92	0,87	11,77	0,88	0,94
Lagos-Nigeria	115	1	2,80	1,18	5,00	1,00	0,94
Macau	597	3	1,91	0,86	6,67	0,33	0,85
Norvège	2 633	3	1,22	0,37	11,67	0,00	0,94
Nouvelle-Zélande	668	1	1,70	0,88	15,00	0,00	0,97
Pakistan	5 228	15	2,90	0,88	12,67	0,83	0,88
Philippines	950	5	2,36	1,07	13,00	0,75	0,95
Portugal	1 191	3	1,73	0,72	15,00	1,00	0,90
Roumanie	281	1	1,13	0,26	15,00	1,00	0,85
Suède	101	1	1,38	0,51	15,00	0,00	0,87
Taiwan	3 129	9	2,05	0,92	10,00	0,67	0,94
Trinidad	228	1	4,20	0,70	15,00	0,00	0,85
Turquie	644	1	1,25	0,55	5,00	0,00	0,88
Vietnam	492	1	3,37	0,34	15,00	0,00	----

Non spécifié	3 600	15	1,64	0,80	15,00	0,31	0,93
-----------------	-------	----	------	------	-------	------	------

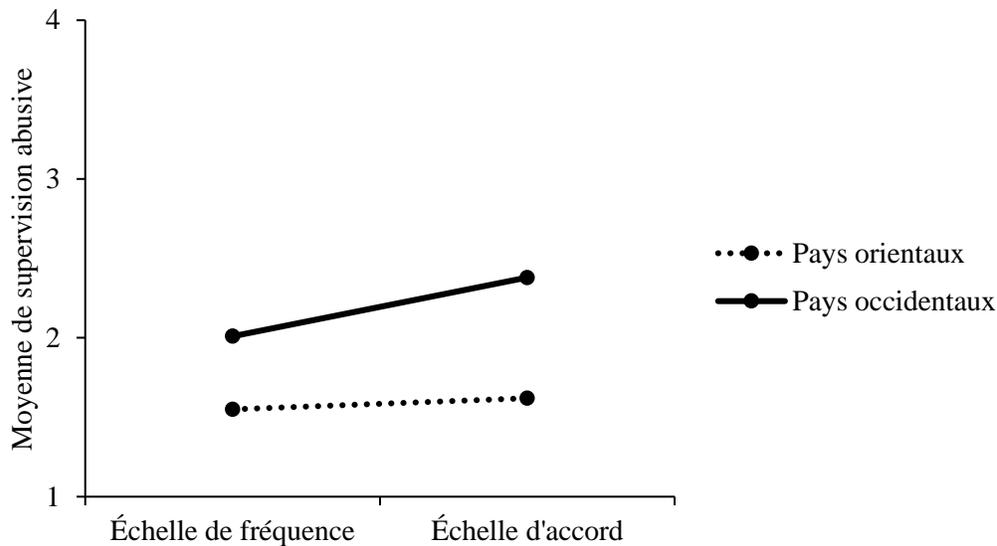
Note. $k = 242$. Le type d'échelle a été codé par 0 pour les échelles de fréquence et par 1 pour les échelles d'accord. N = nombre total de répondants parmi les études; k = nombre d'échantillons; μSA = moyenne agrégée de supervision abusive des échantillons; σSA = moyenne de l'écart-type des données de supervision abusive des échantillons; r_{xx} = moyenne du coefficient de fidélité de l'outil (α) rapporté dans les études; SA = supervision abusive.

En ne tenant compte que des données valides et complètes ($k = 187$), on constate que les pays occidentaux ont davantage tendance à opter pour une échelle de fréquence ($k = 54$, soit 70 % des études occidentales), et que les pays orientaux sont davantage portés à utiliser une échelle d'accord ($k = 62$, soit 56 % des études orientales); l'échelle d'accord étant moins sujette à la variabilité des interprétations d'une culture à l'autre (Vogel et al., 2015). En ce sens, on peut se questionner quant à la différence entre les moyennes observée entre l'Orient et l'Occident : est-elle toujours significative lorsqu'on considère également le type d'échelle utilisé? Une ANOVA factorielle (Abbott, 2016) fut donc réalisée afin de déterminer si la moyenne de supervision abusive générée par l'outil de Tepper (2000) est influencée par l'appartenance à une région du globe (effet principal), par le type d'échelle utilisé (effet principal), ou par l'interaction entre les deux variables (effet d'interaction).

L'ANOVA indique que le modèle généré est significatif ($\alpha = 0,05$, $F(3, 183) = 18,95$, $p < 0,001$) et que sa taille d'effet est grande ($R^2 = 0,24$; Cohen, 1988). L'effet principal de l'appartenance à une région (sur la moyenne de supervision abusive) est significatif ($F(1, 183) = 37,30$, $p < 0,001$) et sa taille d'effet est grande ($\eta^2 = 0,17$). L'effet principal du type d'échelle est aussi significatif ($F(1,183) = 5,01$, $p < 0,05$) et sa taille d'effet est petite ($\eta^2 = 0,03$). L'effet d'interaction entre l'appartenance à une région du globe et le type d'échelle utilisé est toutefois non significatif ($F(1,183) = 2,27$, $p = 0,13$). Une représentation visuelle des données est disponible à la Figure 3.

Figure 3

Moyenne de supervision abusive obtenue par l'outil de Tepper (2000) selon le type d'échelle utilisé ainsi que l'appartenance à une région du globe (Orient c. Occident)



Les constats susmentionnés montrent que le type d'échelle utilisé et l'appartenance à une région du globe auraient tous les deux un effet distinct significatif sur la moyenne de supervision abusive rapportée dans les écrits. Les moyennes des pays orientaux sont plus élevées que celles des pays occidentaux, et la taille d'effet de ce phénomène serait grande. L'effet du type d'échelle a été précédemment discuté à l'objectif 1.

Comparaisons ethniques. La présente recension des écrits a permis d'identifier certains articles qui font des comparaisons entre les résultats obtenus parmi diverses cultures ethniques dans un même devis de recherche. Vogel et al. (2015) ont trouvé que la moyenne de supervision abusive était significativement plus élevée dans la culture asiatique (*Confucian Asian culture*) que dans la culture anglo-saxonne. Les auteurs suggèrent que de tels comportements seraient plus normatifs dans la culture asiatique, et donc que les superviseurs seraient davantage portés à y avoir recours. Dans cette même étude, les auteurs ont trouvé un soutien empirique comme quoi la culture aurait un effet modérateur sur la relation entre la supervision abusive et la perception de justice interpersonnelle, par l'entremise de son influence sur l'orientation envers la hiérarchie des employés (*power distance orientation*).

Haar, de Fluiter et Brougham (2016) ont comparé les résultats obtenus à l'aide de la mesure de Tepper (2000) entre diverses populations œuvrant en Nouvelle-Zélande (cols bleus néo-zélandais, personnes indigènes Maori et travailleurs immigrants chinois). Les employés chinois ont rapporté une perception de supervision abusive plus grande que les deux autres groupes (indices statistiques non présentés dans l'article). Au Canada, aucune comparaison de ce genre n'aurait été effectuée.

Certains auteurs ont reproduit l'étude de leur modèle dans divers pays (chaque pays correspondant à un devis en soi) : Kluemper et al. (2019) ont étudié leur modèle aux États-Unis et en Roumanie; et Lee, Wang et Piccolo (2018) ont testé le leur aux États-Unis et en Corée du Sud. Quoique ces auteurs affirment que leur modèle est généralisable à plus d'une culture, cette façon de procéder ne teste pas l'effet de la culture au moyen d'un test d'hypothèse.

Comparaisons organisationnelles. La majorité des études recensées administre le questionnaire de supervision abusive dans une variété de domaines d'emploi au sein d'un même devis, par exemple en utilisant une plateforme de sondage ou en utilisant le réseau de contacts d'étudiants universitaires. En ce sens, près de 60 % des études ($k = 166$) ont un échantillon de participants non restreint à un secteur d'emploi.

Très peu d'études comparent comment se comporte la mesure dans diverses cultures organisationnelles. Tzafir, Gur et Blumen (2015) ont comparé les moyennes de supervision abusive rapportées par les secteurs privé, public et tertiaire (ANOVA à un facteur, $F = 3,43$, $p < 0,05$). Les comparaisons multiples indiquent que la moyenne des comportements abusifs du secteur privé est supérieure à celle du secteur tertiaire ($p < 0,05$), mais que leur moyenne ne diffère pas significativement de celle du secteur public. En France, Bortolon et al. (2019) rapportent que les comportements abusifs seraient plus fréquents au sein des organisations privées en comparaison aux organisations publiques (*Mann-Whitney test*, $p = 0,003$). En outre, Lam, Walter et Huang (2017) ont testé leur modèle théorique au sein d'un contexte manufacturier (étude 1), et ensuite, dans un contexte de service à la clientèle (étude 2), afin de corroborer leurs résultats dans plus d'un contexte organisationnel (Wright et Sweeney, 2016). Quoiqu'informative, ce type de comparaison ne permet pas d'examiner l'effet du milieu organisationnel à l'aide d'un test d'hypothèse.

Ainsi, les constats présentés ci-dessus indiquent que les moyennes de supervision abusive sont plus élevées en Orient qu'en Occident, et qu'elles pourraient varier selon le contexte organisationnel étudié. Les études qui testent empiriquement l'effet de la culture au sein d'un devis de recherche demeurent rares. Cette différence observée entre les divers employés (ou entre les diverses cultures) est-elle réelle ou plutôt due à la mesure qui se comporte différemment tout dépendant de la population à qui on l'administre? Ceci reste à explorer.

Objectif 7 : l'objectif 7 était de repérer toute étude effectuée sur le leadership tyrannique réalisée dans le milieu de l'éducation.

Le seul outil répertorié ayant été administrée au sein du milieu de l'éducation est l'échelle de Thoroughgood et al. (2012), utilisée par Woestman et Wasonga (2018) aux États-Unis. Leur étude portait

sur le leadership destructeur et non le leadership tyrannique, mais les constats propres à la sous-échelle des comportements dirigés envers les subordonnés sont rapportés. Il s'agit d'ailleurs de la sous-échelle dont l'incidence des comportements sont les plus fréquents. Les données de la sous-échelle étaient significativement corrélées à la satisfaction au travail ($r = -0,46, p < 0,01$), au stress ($r = -0,31, p < 0,05$) et à l'intention de quitter son emploi ($r = -0,45, p < 0,01$) des répondants qui subissent les actes destructeurs.

Discussion

Le présent essai visait à effectuer une revue systématique de la littérature permettant de repérer les outils de mesure du leadership tyrannique utilisés dans les derniers cinq ans et d'effectuer un bilan-synthèse des informations extraites des écrits recensés afin de répondre à sept objectifs de recherche. Les constats obtenus à l'égard des sept objectifs de recherche guident la discussion qui suit portant sur la qualité méthodologique des outils recensés ainsi que les recommandations qui en découlent. Le but ultime est de proposer des recommandations aux chercheurs et aux praticiens œuvrant au sein du milieu scolaire québécois et qui s'intéressent au leadership tyrannique afin de faciliter les choix quant à leur outil de mesure. L'analyse et les recommandations sont regroupées par enjeu de mesure.

Constats, pistes de réflexion et recommandations sur les outils à partir des enjeux identifiés

Multiplicité de versions

Les corrélations effectuées pour l'objectif 1 à l'égard de l'outil de Tepper (2000), tout comme celles obtenues par Mackey et al. en 2017, indiquent que le choix du type d'échelle pourrait avoir un impact sur les résultats obtenus dans les études empiriques, comme quoi l'utilisation d'une échelle de fréquence pourrait générer des moyennes inférieures à celles obtenues à l'aide d'une échelle d'accord. L'ANOVA factorielle effectuée précise que cette différence est significative lorsqu'on a recours à la version originale de l'outil de Tepper (2000) à quinze items, mais pas lorsqu'on opte pour la version courte à cinq items. L'influence du type d'échelle pourrait être examinée davantage, par exemple, en demandant à des participants d'utiliser les deux types d'échelle (accord et fréquence) dans un même questionnaire et de comparer la distribution des données de supervision abusive obtenues par les deux échelles. Il serait intéressant de vérifier si certains items sont davantage influencés par le type d'échelle utilisé que d'autres. Aucun article ayant fait cet exercice n'a été répertorié. On pourrait aussi explorer si les résultats obtenus à l'aide des différents types d'échelle expliquent mieux (ou différemment) la variance de certaines variables du réseau nomologique.

Par ailleurs, les analyses corrélationnelles montrent que le nombre d'items utilisés ne semble pas lié aux résultats de supervision abusive obtenus dans les écrits. Similairement à Breevhart et de Vries (2017; cités dans l'*Introduction*), Garcia et al. (2015) ont obtenu une corrélation de l'ordre de la

singularité ($r = 0,97$, $p < 0,01$) entre la version originale de Tepper (2000) à quinze items et celle de Mitchell et Ambrose (2007) à cinq items. Dans la même veine, Kacmar et al. (2015) ont calculé une corrélation de 0,96 entre la version originale à quinze items et une version raccourcie à six items. Ces résultats semblent à première vue indiquer qu'il serait acceptable d'utiliser une version plus courte.

Toutefois, tel que décrit précédemment, les constats des études de Breevhart et de Vries (2017) et de Pundt et Schwarzbeck (2018) ajoutent un bémol non négligeable, comme quoi la validité du construit pourrait être affectée par le nombre d'items utilisés, mais les deux études rapportent des constats inverses. Ceci pourrait potentiellement expliquer pourquoi la corrélation entre le nombre d'items et les résultats de supervision abusive rapportés dans les écrits recensés par le présent essai était de faible envergure. Les facteurs pouvant expliquer les divers constats obtenus méritent donc d'être explorés davantage.

Mitchell et Ambrose (2007) ont effectué une analyse factorielle exploratoire et indiquent que la version à cinq items (mesurant les comportements actifs) explique 9 % de la variance du construit de supervision abusive, ce qui est très faible (Hair et al., 1995; Hinkin, 1998; Pett et al., 2003). Ce même pourcentage se chiffre à 48 % pour les cinq items mesurant les comportements passifs. Se peut-il que la portion de variance non-expliquée par la version à cinq items (des comportements actifs) soit trop grande pour que l'outil soit suffisamment sensible pour distinguer une différence entre les résultats obtenus par les différents types d'échelle? Le pourcentage de variance expliquée par la version originale à quinze items n'a pas été trouvé. Une avenue de recherche possible serait d'effectuer une analyse factorielle exploratoire avec les deux versions de l'outil et d'en comparer le pourcentage de variance expliquée. D'autres propriétés psychométriques pourraient en même temps être vérifiées, notamment la fidélité test-retest.

Même si la présente recension a permis d'identifier quelques auteurs qui ont comparé les résultats de la version originale de l'outil de Tepper (2000) à ceux d'une variante plus courte, le nombre d'études reste petit et cela pourrait être reproduit. La validité de construit et les propriétés psychométriques sont-elles véritablement affectées par le nombre d'items utilisé? Cela reste à explorer.

Recommandation 1 : piloter le questionnaire qu'on souhaite utiliser au sein du milieu de l'éducation, en demandant aux participants de compléter les mêmes questions à l'aide d'une échelle d'accord et de fréquence. Vérifier en quoi les données générées par les deux échelles diffèrent. Examiner, aussi, si les résultats calculés sont différents advenant qu'on ne tienne compte que d'un nombre réduit d'items de l'échelle (soit ceux qu'on juge plus pertinents pour la clientèle étudiée).

Multiplicité des conceptualisations

Même si la vaste majorité des études quantitatives effectuées au cours des cinq dernières années utilisent l'échelle de Tepper (2000) ou une variante de celle-ci, la présente recension a mis en lumière plusieurs autres outils utilisés pour mesurer le leadership tyrannique et, tel qu'attendu, ceux-ci divergent quant au nombre de dimensions à attribuer au construit. Peu d'études recensées s'intéressent à explorer en quoi cette diversité de façons de concevoir le leadership tyrannique permet de mieux prédire (ou de prédire différemment) le réseau nomologique de la variable. Cela reste à vérifier.

Recommandation 2 : mesurer le leadership tyrannique au sein du milieu de l'éducation à l'aide de deux outils de mesure et en vérifier la validité incrémentielle. L'outil de Tepper (2000) devrait être priorisé étant donné que les comparaisons avec les constats publiés dans les écrits scientifiques en seraient facilitées (Lavoie-Tremblay et al., 2016 pour la version française). Les échelles de Cacciatore (2015) et d'Aasland et al. (2010) ayant été traduites en français et utilisées au Québec (Boudrias et al., 2020; Trépanier et al. 2019), elles pourraient être incluses comme deuxième mesure.

Diversité de sources

Même si la majorité des outils sont strictement utilisés afin de mesurer la supervision abusive d'un point de vue centré sur la victime (complétés par les subordonnés), plusieurs études récentes ont néanmoins adopté une perspective centrée sur le leader (données auto-rapportées par les leaders) en ayant recours à l'échelle de Tepper (2000) ou de Johnson et al. (2012). Ceci rejoint une recommandation de Mackey et al. (2017) qui suggèrent que la perspective des leaders à l'égard de leurs propres comportements peut être informative. Certains chercheurs ont également modifié l'outil de Tepper (2000) afin d'opter pour une perspective centrée sur une tierce partie observatrice (données rapportées par les collègues). Le choix de perspective doit être conséquent avec les autres variables incluses dans les modèles testés (ex. : Yeung et Shen, 2019). Toutefois, peu d'études explorent en quoi les données provenant de diverses sources (ex : leader, subordonnés, collègues) sont comparables les unes aux autres ainsi que les facteurs influençant ces corrélations.

Recommandation 3 : déterminer qui complètera l'échelle de supervision abusive (subordonnés, leader, collègues) en fonction des variables (du réseau nomologique) étudiées (variables d'intérêt du modèle testé). Administrer le même questionnaire à diverses sources et tester en quoi les résultats générés par l'outil diffèrent selon que la supervision abusive d'un leader envers un(e) subordonné(e) soit évaluée par le leader, le(la) subordonné(e) ou un collègue.

Réurrence et prévalence

La façon idéale de capter la récurrence des manifestations abusives (et leurs variations au fil du temps) serait d'avoir recours à un devis longitudinal, incluant des journaux de bord que les participants complèteraient à plusieurs reprises (ex. : Barnes et al., 2015; Bormann, 2017; Johnson et al., 2012). À cet égard, au cours des dernières années, on observe une tendance à opter pour des mesures de leadership « au jour-le-jour » (ex. : Ju et al., 2019). Kelemen, Matthews et Breevhart (2020) ont produit une méta-analyse sur le sujet. Quoique leur utilisation soit à la hausse, les devis longitudinaux demeurent peu utilisés, limitant ainsi la capacité à identifier des liens de causalité (Martinko et al., 2013). Les études ayant recours à un devis longitudinal seraient également inexistantes dans le milieu de l'éducation.

Recommandation 4a : si possible, opter pour des devis longitudinaux.

En revanche, lorsque les chercheurs optent pour un devis transversal, il n'est pas clair quel type d'échelle serait préférable, chacun ayant ses avantages et ses limites. Les échelles de fréquences ont été utilisées pour estimer la prévalence du phénomène au sein de populations (Brunet et al., 2015; Nielsen et al., 2016). Elles suivent habituellement la prémisse que plus un comportement est fréquent, plus l'intensité de l'abus est grande. Les échelles d'accord pourraient être priorisées dans un contexte où on souhaite étudier plusieurs populations ou cultures, étant moins sujettes à la variabilité des interprétations (Vogel et al., 2015). Cela dit, peu importe le type d'échelle utilisé dans un devis transversal, il ne permettrait pas de saisir complètement l'aspect de récurrence des comportements abusifs, tel qu'il est compris dans les définitions du construit (ex. : en ne tenant pas compte de certains comportements moins fréquents mais, qui peuvent occasionner de la souffrance), ni d'analyser les nuances ou l'évolution des manifestations comportementales au fil du temps.

D'un point de vue empirique, on pourrait examiner si un type d'échelle engendrerait des résultats qui sont plus près de la « vraie valeur », c'est-à-dire de la prévalence réelle du phénomène au sein du milieu étudié. Tel que le montre les résultats de la présente recension des écrits, la supervision abusive peut être vue comme un phénomène peu fréquent (Mackey et al., 2017). La moyenne des résultats obtenus à l'aide des échelles de fréquence, étant inférieure à celle générée par les échelles d'accord, serait-elle plus près de la valeur réelle?

Cela dit, la présente recension montre que peu d'études récentes estiment la prévalence du phénomène. Le seuil de fréquence à utiliser pour déterminer la prévalence demeure un enjeu non résolu. Breevhart et de Vries (2017) vont jusqu'à remettre en question si la supervision abusive peut véritablement s'avérer être un phénomène peu fréquent. Par exemple, serait-ce plutôt dû aux mesures qui ne captent pas bien le phénomène? Serait-ce un effet dû au recrutement (ex. : les leaders ou les employés

qui participent aux études ont-ils tendance à être plus agréables et humbles et donc moins portés à être abusifs)? Breevhart et de Vries (2017) mentionnent aussi le manque d'écrits empiriques testant le fait que la supervision abusive puisse être plus commune dans certains types d'industries. En étudiant divers milieux organisationnels, nous pourrions ainsi avoir une meilleure idée de la prévalence du phénomène. Une avenue de recherche future pourrait être d'en estimer la prévalence et de vérifier si un type d'échelle engendre des résultats plus près de cette valeur.

Recommandation 4b : estimer la prévalence du phénomène au sein des établissements scolaires au Québec en utilisant une échelle de fréquence, et rapporter les résultats en incluant la prévalence estimée selon les divers seuils (ex. : prévalence obtenue si on tient compte des comportements qui surviennent occasionnellement ainsi que celle obtenue en ne tenant compte que des comportements fréquents).

Exhaustivité et perception de victimisation

L'exhaustivité des outils est peu abordée dans les écrits récents. Tel que mentionné précédemment, les auteurs ont d'ailleurs tendance à privilégier l'utilisation de questionnaires courts plutôt qu'exhaustifs (Cacciatore, 2015; Peng et al., 2014). Pourtant, Blase, Blase et Du (2008) ont listé les comportements de mauvais traitement (*mistreatment*) les plus fréquemment manifestés envers des enseignant(e)s américain(e)s ainsi que ceux engendrant la plus grande intensité de souffrance. Lorsqu'on compare les quinze items de Tepper (2000) avec le top quinze des deux listes, on constate que l'outil de Tepper (2000) ne serait pas exhaustif (Tableau 6). Le « mauvais traitement » étant susceptible d'inclure des comportements moins hostiles que la supervision abusive, cela est en quelque sorte normal, mais il reste que certains des comportements qui engendreraient le plus de souffrance (ex. : *gave me unwarranted reprimands* et *unjustly criticized me*) sont tout de même couverts dans l'outil de Cacciatore (2015), mais pas dans celui de Tepper (2000). Il est donc important de s'assurer que les comportements abusifs centraux au construit (ayant une valeur prédictive dans le milieu étudié) soient inclus dans l'instrument utilisé.

Tableau 6

Comparaison des items de l'outil de Tepper (2000) avec les manifestations abusives les plus fréquentes ainsi que celles occasionnant la plus grande intensité de souffrance identifiées par une population d'enseignants dans l'article de Blase, Blase et Du (2008)

Tepper (2000)	Comportements les plus fréquents	Comportements engendrant la plus grande intensité de souffrance
1. Ridicules me	1. Failed to recognize or praise me	1. Tried to intimidate me
2. Tells me my thoughts or feelings are stupid	for work-related achievements	2. Failed to recognize or praise me for work-related
3. Gives me the silent treatment	2. Favored "select" teachers*	achievements
	3. Tried to intimidate me	

4. Puts me down in front of others	4. Failed to support me (i.e. abandoned me) in difficult interactions with students and/or parents*	3. Failed to support me (i.e. abandoned me) in difficult interactions with students and/or parents*
5. Invades my privacy	5. Ignored or snubbed me	4. Gave me unwarranted reprimands*
6. Reminds me of my past mistakes and failures	6. Nitpicked about time or micromanaged me*	5. Made unreasonable demands*
7. Doesn't give me credit for jobs requiring a lot of effort	7. Was insensitive to my personal matters	6. Favored "select" teachers*
8. Blames me to save himself/herself embarrassment	8. Made unreasonable demands*	7. Lied to me or about me
9. Breaks promises he/she makes	9. Stonewalled or failed to respond to me	8. Nitpicked about time or micromanaged me*
10. Expresses anger at me when he/she is mad for another reason	10. Used a "snitch's" (other teacher or student) information to criticize me*	9. Used negative terms to label me and my behavior*
11. Makes negative comments about me to others	11. Overloaded me with work and/or overloaded me with trivial work*	10. Unjustly criticized me*
12. Is rude to me	12. Lied to me or about me	11. Used a "snitch's" (other teacher or student) information to criticize me*
13. Does not allow me to interact with my coworkers	13. Attempted to isolate me or control my interactions with other teachers and colleagues	12. Stonewalled or failed to respond to me
14. Tells me I'm incompetent	14. Spied on me*	13. Spied on me*
15. Lies to me	15. Used negative terms to label me and my behavior*	14. Overloaded me with work and/or overloaded me with trivial work*
		15. Was insensitive to my personal matters

Note. *Item qui n'est pas couvert par l'outil de Tepper (2000).

Lorsque possible, Sasso et Gonzalez-Morales (2018) suggèrent de bonifier les mesures actuelles en permettant aux répondants de s'exprimer sur d'autres comportements manifestés par leur superviseur. La nature qualitative de ce type de réponse pourrait renseigner les chercheurs sur les comportements qui pourraient éventuellement être inclus dans la mesure, tout dépendant de la clientèle et du contexte visé par l'étude. Aucune étude n'a été repérée ayant appliqué cette recommandation.

Dans toutes les études recensées, les instruments de mesure sont utilisés de manière à calculer un score moyen global, traitant ainsi chacun des items comme des comportements équivalents les uns aux autres. Tel que mentionné précédemment, ceci peut être problématique, dans la mesure où chaque comportement peut engendrer une intensité de souffrance différente (Blase et al., 2008), et aussi dans l'optique où certaines cultures organisationnelles ou ethniques pourraient différer sur ce qu'ils considèrent approprié (Tepper, et al., 2017; Vogel et al., 2016; Zhang et Liu, 2018). Afin de mieux tenir compte de la perspective des répondants, Sasso et Gonzalez-Morales (2018) suggèrent le recours à des mesures réflexives, permettant aux participants d'élaborer à l'égard des facteurs contextuels et de leur perception de victimisation. En ce sens, Hershcovis (2011) indique que les études de type journaux de bord (*event-based diary study*) et l'approche des incidents critiques pourraient enrichir notre compréhension du

phénomène. Tel que mentionné précédemment, quoique ce type de méthode fut utilisé par quelques études recensées, on pourrait davantage y avoir recours.

Similairement, Tepper et al. (2017) discutent de la perspective phénoménologique des présumées victimes d'abus. Ils s'interrogent sur la variation de la perception d'abus selon les circonstances organisationnelles (ex : quelles en sont les manifestations dans un contexte militaire?) et se demandent à quel point la perception de victimisation est réellement essentielle pour ressentir les conséquences de l'abus (ex : est-ce que les personnes qui ne se considèrent pas victimes de supervision abusive en ressentent aussi les conséquences?). Tepper et al. (2015) traitent cette perspective de victimisation comme une variable médiatrice. Ceci dit, devrait-on constamment inclure une mesure de la perception de victimisation aux devis de recherche? Devrait-on plutôt en tenir compte dans le choix des items? Les réponses à ces questions offriraient l'occasion d'ajuster les mesures au besoin.

En outre, une autre avenue de recherche possible serait d'examiner à quel point la fréquence versus l'intensité perçue des comportements abusifs permet de mieux prédire les liens entre le construit et ses conséquences (Hershcovis, 2011).

Recommandation 5 : en utilisant un devis qualitatif, répéter ce que Blase et al. (2008) ont fait aux États-Unis au sein du système d'éducation québécois, afin d'identifier les comportements abusifs des leaders qui sont les plus fréquents et ceux générant une plus grande souffrance auprès des victimes, de sorte à pouvoir adapter, au besoin, les outils de mesure existants. Un tel devis qualitatif permettrait de discuter de la perspective phénoménologique des victimes d'abus tel que proposé par Tepper et al. (2017). De plus, lorsque possible, bonifier les mesures actuelles en permettant aux répondants de s'exprimer sur d'autres comportements manifestés par leur superviseur.

En outre, les subordonnés qui sont témoins de manifestations abusives de la part de leur superviseur envers certains de leur collègues (sans que les manifestations leur soient destinées directement) peuvent aussi ressentir un mal-être (Priesemuth et Schminke, 2019). C'est d'ailleurs pour tenir compte de la perspective et des impacts ressentis par les témoins de la supervision abusive que certains chercheurs ont maintenant recours à des instruments mesurant la supervision abusive vicariante, tel que mentionné précédemment à l'objectif 3 (ex. : Huang et al., 2019; Shao et al. 2018; Zhang et al., 2019). Les recommandations de Sasso et Gonzalez-Morales (2018) et d'Herschcovis (2011) pourraient tout autant être utiles pour enrichir notre compréhension de la perspective centrée sur une tierce partie.

Comparaisons organisationnelles et ethniques

Les résultats du présent essai montrent que la supervision abusive est maintenant étudiée dans des dizaines de pays. Toutefois, peu d'études comparent la manière dont se comportent les instruments de

mesures au sein de divers cultures organisationnelles et ethniques. L'étude d'Hu, Wu et Wang (2011) est un des rares exemples où on a comparé la dimensionnalité du construit entre deux cultures ethniques (américaine et taiwanaise dans leur cas). Ils ont trouvé que la conceptualisation à un facteur se généralisait aux deux cultures, mais que certains items pourraient être perçus différemment par les deux cultures (certains items ayant montré de la *non-invariance* d'une culture à l'autre). Les auteurs élaborent des pistes d'exploration pouvant influencer comment les deux populations conceptualisent la supervision abusive.

Il serait d'ailleurs intéressant d'explorer la validité des mesures actuelles parmi les minorités visibles d'un même pays (d'autres variables démographiques ayant davantage fait l'objet d'études, dont le sexe et l'âge des subordonnés; voir Wang et al., 2019 pour une méta-analyse sur le sujet). Le Québec ayant une grande diversité culturelle, ce type de comparaison pourrait même être fait au sein du système d'éducation québécois (ex. : minorités visibles, autochtones, francophones, anglophones, etc.). Dans une étude au Canada, Frone et Blais (2019) comparent la prévalence de la fatigue physique, mentale et émotionnelle entre diverses professions au sein des Forces de l'air. La même expérience pourrait être réalisée pour la supervision abusive entre divers milieux organisationnels ou ethniques.

Cela dit, comment savoir si les différences de moyennes observées entre les divers milieux sont réelles ou simplement un effet de la mesure qui soit mal adaptée aux contextes ou populations étudiés? Les recommandations effectuées par Sasso et Gonzales-Morales (2018), Hershovic (2011) et Tepper et al. (2015) mentionnées précédemment pourraient être un bon point de départ pour explorer les différences de manifestations abusives et la perception de victimisation au sein des écoles québécoises. Les analyses quantitatives réalisées par Hu et al. (2011) pourraient également fournir des pistes de réflexion intéressantes quant à la façon dont les divers milieux québécois conceptualisent le construit.

Somme toute...

L'objectif 7 montre à quel point peu d'études sont réalisées dans le milieu d'éducation, et encore moins au sein du système scolaire québécois. Le présent travail visait à effectuer un certain bilan des tendances actuelles quant aux choix qui sont faits et qui peuvent influencer la qualité méthodologique des outils. À cet égard, certaines recommandations ont pu être identifiées. Celles-ci se résument, d'abord et avant tout, à valider l'outil qu'on veut utiliser dans le milieu scolaire. Le recours à un devis qualitatif permettrait d'explorer la perspective phénoménologique des présumées victimes au sein de la population cible, de vérifier l'exhaustivité et la pertinence des items et d'apporter les ajustements nécessaires à la mesure, au besoin. Le recours à une échelle de fréquence permettrait d'estimer la prévalence du phénomène, mais il serait intéressant d'explorer en quoi le choix du type d'échelle (et d'un nombre réduit du nombre d'items) influence les résultats générés dans le cadre du projet-pilote de l'outil.

Limites de l'étude

La présente recension comporte toutefois des limites non négligeables. Premièrement, seuls certains enjeux de mesure ont été discutés. Il est possible que d'autres enjeux (ex. : la désirabilité sociale des items) puissent influencer la qualité des instruments. Deuxièmement, il se pourrait que la démarche utilisée n'ait pas permis de repérer tous les articles publiés sur la supervision abusive au cours des derniers cinq ans. Peut-être que certains constats, qui auraient influencé les recommandations, n'ont pas été rapportés. Cela dit, le recours à cinq bases de données et l'utilisation d'une combinaison de mots-clés permettant de ratisser plus largement les publications (ex. : permettant de cibler les articles portant sur le leadership destructeur, qui est un construit plus large que leadership tyrannique) furent un effort pour contrer cette limite. Troisièmement, l'extraction et la compilation des informations des 270 articles recensés a été réalisée par une seule personne. Même si une vérification aléatoire fut réalisée, il se peut que des erreurs se soient glissées. Seules les informations disponibles dans les articles ont été utilisées. L'auteure n'a pas cherché à communiquer avec les auteurs lorsque des informations étaient manquantes. Quatrièmement, le contexte de pandémie en cours pourrait influencer la validité des mesures actuelles, étant donné que la supervision abusive serait susceptible de se manifester différemment, par exemple, dans les milieux effectuant davantage de télétravail (ex. : Walsh et al., 2019).

Conclusion

D'un point de vue plus général, les enjeux liés à la mesure du leadership tyrannique pourraient remettre en question, ou du moins nuancer, les conclusions des écrits empiriques. De par nos choix méthodologiques actuels, ignorons-nous certaines manifestations abusives qui continuent de dévaster des victimes et des organisations? Établissons-nous des liens corrélationnels ou causaux sans que les comportements énumérés soient considérés comme néfastes par les présumées victimes? Les multiples adaptations des instruments de mesure affectent-elles les résultats obtenus? En quoi les outils sont adaptés (ou pas) aux diverses cultures organisationnelles et ethniques? Le leadership tyrannique est-il réellement un phénomène peu fréquent? Ces questions sont pertinentes pour tous les milieux d'échantillonnage et le système d'éducation québécois n'en fait pas exception.

Le besoin de mieux étudier les enjeux liés à la mesure du leadership tyrannique se fait ressentir. Les constats et les recommandations qui en découleront permettront de mieux appuyer les chercheurs et les praticiens qui souhaitent effectuer des choix éclairés concernant leurs outils et leurs devis de recherche.

En ce qui concerne spécifiquement le système scolaire québécois, les chercheurs et les praticiens possèdent-ils un instrument mesurant la supervision abusive qui y est adapté et validé? La réponse est non. En ce sens, et à la lumière des constats présentés dans le présent essai, voici les étapes qui pourraient être

réalisées : (1) **identifier les manifestations de supervision abusive** les plus fréquentes ainsi que celles engendrant la plus grande intensité de souffrance auprès des enseignant(e)s québécois(e)s en reproduisant l'étude de Blase et al. (2008); (2) **effectuer une analyse comparative entre les outils** de mesure préexistants, qui ont été répertoriés dans la présente étude, et les deux listes de comportements (générées à l'étape 1), afin d'identifier quelles échelles et sous-échelles chevauchent le plus les comportements abusifs les plus fréquents et engendrant la plus grande intensité de souffrance (une analyse conceptuelle à la lumière des théories du leadership pourrait aussi être incluse à cette étape); et (3) **valider empiriquement les choix effectués** par rapport à l'outil de mesure. La validation empirique pourrait servir à déterminer : (3a) la validité incrémentielle d'une échelle ou d'une sous-échelle en rapport à une autre (ex. : est-ce que l'ajout de la sous-dimension des punitions non-contingentes de Cacciatore (2015) à l'outil de Tepper (2000) permet d'expliquer une plus grande portion de la variance de certaines variables d'intérêt) (objectif 2); (3b) advenant l'utilisation de l'échelle de Tepper (2000), en quoi les résultats générés à l'aide d'une version courte sont similaires à ceux générés avec la version originale (ex. : sont-ils corrélés de l'ordre de la singularité? Est-il approprié d'utiliser moins d'items?) (objectif 1); (3c) en quoi les résultats générés diffèrent selon le type d'échelle utilisé, et qu'en est-il de la prévalence advenant qu'on utilise une échelle plutôt qu'une autre (objectif 1); (3d) en quoi l'outil génère des résultats différents parmi les diverses communautés culturelles (géographiques, ethniques) du Québec et explorer pourquoi on observe ces différences (objectif 6); (3e) en quoi les résultats diffèrent si on administre le questionnaire à plus d'une source (objectif 3); (3f) en quoi les manifestations de supervision abusive varient d'une journée à l'autre ou au fil de temps (objectif 4).

Étant donné que le système d'éducation québécois est un tout nouveau terrain de collecte pour l'étude de la supervision abusive, on gagnerait à effectuer ces trois grandes étapes. Un tel outil adapté et validé au sein du milieu scolaire québécois pourrait servir à long terme afin d'identifier les établissements ou les commissions scolaires qui nécessitent une attention particulière et de suivre l'évolution de la prévalence du phénomène. Au fil du temps, une meilleure compréhension des antécédents et des mécanismes qui promeuvent un tel leadership au sein du milieu de l'éducation permettraient d'adapter les pratiques organisationnelles en matière de ressources humaines afin de contrer les effets néfastes associés au leadership tyrannique.

Références

- Aasland, M. S., Skogstad, A., Notelaers, G., Nielsen, M. B. et Einarsen, S. (2010). The prevalence of destructive leadership behaviour. *British Journal of Management*, 21(2), 438-452.
doi:10.1111/j.1467-8551.2009.00672.x
- Abbott, M. L. (2016). *Using statistics in the social and health sciences with SPSS® and Excel®*. John Wiley & Sons, Inc.
- Akhtar, S. et Shaukat, K.(2016). Impact of petty tyranny on alienation from work: Role of self-esteem and power-distance. *Global Journal of Flexible Systems Management*, 17(3), 275-285. doi :
10.1007/s40171-016-0126-4
- Ambrose, M. L. et Ganegoda, D. B. (2020). Abusive according to whom? Manager and subordinate perceptions of abusive supervision and supervisors' performance. *Journal of Organizational Behavior*, 1-20. <https://doi.org/10.1002/job.2472>
- Aryee, S., Chen, Z. X., Sun, L.-Y. et Debrah, Y. A. (2007). Antecedents and outcomes of abusive supervision: Test of a trickle-down model. *Journal of Applied Psychology*, 92(1), 191-201.
- Ashforth, B. E. (1994). Petty tyranny in organizations. *Human Relations*, 47(7), 755-778.
doi:10.1177/001872679404700701
- Bakker, A. B. et Demerouti, E. (2017). Job demands-resources theory: Taking stock and looking forward. *Journal of occupational health psychology*, 22(3), 273-285. doi:10.1037/ocp0000056
- Barnes, C. M., Lucianetti, L., Bhave, D. P. et Christian, M. S. (2015). “You wouldn’t like me when I’m sleepy” : Leaders’ sleep, daily abusive supervision, and work unit engagement. *Academy of Management Journal*, 58(5),1419-1437. <http://dx.doi.org/10.5465/amj.2013.1063>
- Baumeister, R. F., Bratslavsky, E., Muraven, M. et Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252–1265.
<https://doi.org/10.1037/0022-3514.74.5.1252>
- Berry, C. M., Carpenter, N. C. et Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology*, 97(3), 613–636. doi : 10.1037/a0026739
- Blase, J., Blase, J. et Du, F. (2008). The mistreated teacher: A national study. *Journal of Educational Administration*, 263-301. doi: 10.1108/09578230810869257
- Bormann, K. C. (2017). Linking daily ethical leadership to followers’ daily behaviour: The roles of daily work engagement and previous abusive supervision. *European Journal of Work and Organizational Psychology*, 26(4), 590-600. 10.1080/1359432X.2017.1331217
- Bortolon, C., Lopes, B., Capdevielle, D., Macioce, V. et Raffard, S. (2019). The roles of cognitive avoidance, rumination and negative affect in the association between abusive supervision in the

- workplace and non-clinical paranoia in a sample of workers working in France. *Psychiatry Research*, 271, 581-589. doi : 10.1016/j.psychres.2018.11.065
- Boudrias, J.-S., Rousseau, V. et Lajoie, D. (2020). How lack of integrity and tyrannical leadership of managers influence employee improvement-oriented behaviors. *Journal of Business Ethics*. <https://doi.org/10.1007/s10551-020-04494-5>
- Braun, S. (2017). Leader narcissism and outcomes in organizations: A review at multiple levels of analysis and implications for future research. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.00773
- Breevaart, K. et de Vries, R. E. (2017). Supervisor's HEXACO personality traits and subordinate perceptions of abusive supervision. *Leadership Quarterly*, 28(5), 691-700. <http://dx.doi.org/10.1016/j.leaqua.2017.02.001>
- Bregenzer, A., Felfe, J., Bergner, S. et Jimenez, P. (2019). How followers' emotional stability and cultural value orientations moderate the impact of health-promoting leadership and abusive supervision on health-related resources. *German Journal of Human Resource Management-Zeitschrift Fur Personalforschung*, 33(4), 307-336. doi : 10.1177/2397002218823300
- Brown, M. E. et Mitchell, M S. (2010). Ethical and unethical leadership: Exploring new avenues for future research. *Business Ethics Quarterly*, 20(4), 583-616.
- Brunet, L., Morin, D., Alderson, M., Cacciatore, G., Savoie, A., Boudrias, J.-S. et Nelson, K. (2015). L'incidence du leadership destructeur sur les attitudes et les comportements au travail. *Humain et organisation*, 1(1), 25-29.
- Cacciatore, G. (2015). *Petty tyranny as a form of destructive leadership : Results of studies based on a French-Canadian measure* (Thèse ou essai doctoral accepté, Université du Québec à Montréal, Montréal (Québec, Canada)). Repéré à <https://archipel.uqam.ca/8616/>
- Campbell, W. K. et Campbell, S. M. (2009). On the self-regulatory dynamics created by the peculiar benefits and costs of narcissism: A contextual reinforcement model and examination of leadership. *Self and Identity*, 8, 214-232. doi: 10.1080/15298860802505129
- Chen, S.-C. et Liu, N.-T. (2019). When and how vicarious abusive supervision leads to bystanders' supervisor-directed deviance: A moderated-mediation model. *Personnel Review*, 48(7), 1734-1755. doi : 10.1108/PR-09-2018-0368
- Cheng, B., Chou, L., Wu, T., Huang, M. et Farh, J. (2004). Paternalistic leadership and subordinate responses: Establishing a leadership model in Chinese organizations. *Asian Journal of Social Psychology*, 7, 89-117.
- Choi, W., Kim, S. L. et Yun, S. (2019). A social exchange perspective of abusive supervision and knowledge sharing: Investigating the moderating effects of psychological contract fulfillment and

- self-enhancement motive. *Journal of Business and Psychology*, 34(3), 305-319. doi : 10.1007/s10869-018-9542-0
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^e éd.) Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159. <https://doi.org/10.1037//0033-2909.112.1.155>
- Conseil supérieur de l'éducation. (1994). *Rénover le curriculum du primaire et du secondaire*. Québec, Québec : Les Publications du Québec.
- Corriveau, L. (2004). Identification professionnelle ou suridentification à la profession? La situation de directrices et de directeurs d'établissements scolaires québécois. *Éducation et francophonie*, 32(2), 95-110.
- Cortina, L. M., Rabelo, V. C. et Holland, K. J. (2018). Beyond blaming the victim: Toward a more progressive understanding of workplace mistreatment. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 11(1), 81–100. <https://doi.org/10.1017/iop.2017.54>
- Courtright, S. H., Gardner, R. G., Smith, T. A., McCormick, B. W. et Colbert, A. E. (2016). My family made me do it: A cross-domain, self-regulatory perspective on antecedents to abusive supervision. *Academy of Management Journal*, 59(5), 1630-1652. doi : 10.5465/amj.2013.1009
- Cullen, K. L., Gentry, W. A. et Yammarino, F. J. (2015). Biased self-perception tendencies: Self-enhancement/self-diminishment and leader derailment in individualistic and collectivistic cultures. *Applied Psychology: An International Review*, 64(1), 161-207. doi: 10.1111/apps.12026
- Dionisi, A. M. et Barling, J. (2019). What happens at home does not stay at home: The role of family and romantic partner conflict in destructive leadership. *Stress and Health*, 35(3), 304-317. doi : 10.1002/smi.2858
- Einarsen, S., Aasland, M. S. et Skogstad, A. (2007). Destructive leadership behaviour: A definition and conceptual model. *The Leadership Quarterly*, 18(3), 207-216. doi:10.1016/j.leaqua.2007.03.002
- Epitropaki, O. et Martin, R. (2004). Implicit leadership theories in applied settings: Factor structure, generalizability, and stability over time. *Journal of Applied Psychology*, 89, 293–310. doi :10.1037/0021-9010.89.2.293
- Fatima, N. et Bashir, S. (2017). Exploring supervisor's power distance orientation in service sector: Understanding how it affect employee morale via ego depletion and abusive supervision. *Journal of Managerial Sciences*, 11(3), 529-558.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327–358. <https://doi.org/10.1037/h0061470>
- Forget, M. (2016). *Étude des inducteurs de l'intention de quitter chez une population infirmière québécoise* (Université de Montréal). Repéré à

https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/19050/Forget_Mathieu_2016_these.pdf?sequence=2&isAllowed=y

- Frone, M. R. et Blais, A.-R. (2019). Work Fatigue in a non-deployed military setting: Assessment, prevalence, predictors, and outcomes. *International Journal of Environmental Research and Public Health*, 16(16), 2892. doi : 10.3390/ijerph16162892
- Gabler, C. B., Nagy, K. R. et Hill, R. P. (2014). Causes and consequences of abusive supervision in sales management: A tale of two perspectives. *Psychology and Marketing*, 31(4), 278-293. doi : 10.1002/mar.20694
- Garcia, P. R. J. M., Wang, L., Lu, V., Kiazad, K. et Restubog, S. L. D. (2015). When victims become culprits: The role of subordinates' neuroticism in the relationship between abusive supervision and workplace deviance. *Personality and Individual Differences*, 72, 225-229. doi : 10.1016/j.paid.2014.08.017
- Gatewood, R. D., Field, H. S. et Barrick, M. (2008). *Human Resource Selection*. Thomson South-western.
- Gatti, P., Schyns, B., Cortese, C. G. et Hall, R. J. (2019). *Investigating constructive and destructive leadership styles together: A pattern approach* [communication orale]. 19th European Association of Work and Organizational Psychology Congress: Working for the greater good, Turin, Italie.
- Glaso, L., Skogstad, A., Notelaers, G. et Einarsen, S. (2018). Leadership, affect and outcomes: symmetrical and asymmetrical relationships. *Leadership & Organization Development Journal*, 39(1), 51-65. doi : 10.1108/LODJ-08-2016-0194
- Green, J. E. (2014). Toxic leadership in educational organizations. *Education Leadership Review*, 15(1), 18-33.
- Haar, J. M., de Fluiter, A. et Brougham, D. (2016). Abusive supervision and turnover intentions: The mediating role of perceived organisational support. *Journal of Management & Organization*, 22(2), 139-153. doi : 10.1017/jmo.2015.34
- Hair, J. F., Anderson, R. E., Tatham, R. L. et Black, W. C. (1995). *Multivariate data analysis*. Prentice-Hall.
- Hallinger, P. (2003). Leading educational change: Reflections on the practice of instructional and transformational leadership. *Cambridge Journal of Education*, 33(3), 329-352.
- Harris, K., Harvey, P. et Kacmar, K. (2011). Abusive supervisory reactions to coworker relationship conflict. *The Leadership Quarterly*, 22(5), 1010- 1023.
- Hawass, H. H. (2019). The riddle of self-centered leadership in Arab organizations: A measurement scale. *Management Research Review*, 42(4), 430-459.
- Hershcovis, M. S. (2011). "Incivility, social undermining, bullying...oh my!": a call to reconcile constructs within workplace aggression research. *Journal of Organizational Behavior*, 32, 499-

519. doi: 10.1002/job.689
- Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, 2(1), 104-121.
<https://doi.org/10.1177%2F109442819800100106>
- Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, 44(3), 513-524.
- Hobfoll, S. E., Halbesleben, J., Neveu, J.-P. et Westman, M. (2018). Conservation of resources in the organizational context: The reality of resources and their consequences. *Annual Review of Organizational Psychology and Organizational Behavior*, 5, 103-128.
<https://doi.org/10.1146/annurev-orgpsych-032117-104640>
- Hon, A. H. Y. et Lu, L. (2016). When will the trickle-down effect of abusive supervision be alleviated? The moderating roles of power distance and traditional cultures. *Cornell Hospitality Quarterly*, 57(4), 421-433. doi : 10.1177/1938965515624013
- Hu, C., Wu, T.-Y. et Wang, Y.-H. (2011). Measurement equivalence/invariance of the abusive supervision measure across workers from Taiwan and the United States. *The Journal of Psychology*, 145(2), 111-131. doi : 10.1080/00223980.2010.542505
- Huang, J., Guo, G., Tang, D., Liu, T. et Tan, L. (2019). An eye for an eye? Third parties' silence reactions to peer abusive supervision: The mediating role of workplace anxiety, and the moderating role of core self-evaluation. *International Journal of Environmental Research and Public Health*, 16, 5027. doi :10.3390/ijerph16245027
- Hunter, J. E. et Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.
- Jiang, W. et Gu, Q. (2016). How abusive supervision and abusive supervisory climate influence salesperson creativity and sales team effectiveness in China. *Management Decision*, 54(2), 455-475. doi : 10.1108/MD-07-2015-0302
- Jiang, W., Gu, Q. et Tang, T. L.-P. (2019). Do Victims of Supervisor Bullying Suffer from Poor Creativity? Social Cognitive and Social Comparison Perspectives. *Journal of Business Ethics*, 157(3), 865-884. doi: 10.1007/s10551-017-3660-x
- Jonason, P. K. et Webster, G. D. (2010). The dirty dozen: A concise measure of the dark triad. *Psychological Assessment*, 22(2) (2010), pp. 420-432. <http://dx.doi.org/10.1037/a0019265>
- Johnson, R. E., Venus, M., Lanaj, K., Mao, C. et Chang, C. H. (2012). Leader identity as an antecedent of the frequency and consistency of transformational, consideration, and abusive leadership behaviors. *Journal of Applied Psychology*, 97, 1262-1272. doi : 10.1037/a0029043
- Ju, D., Huang, M., Liu, D., Qin, X., Hu, Q. et Chen, C. (2019). Supervisory consequences of abusive

- supervision: An investigation of sense of power, managerial self-efficacy, and task-oriented leadership behavior. *Organizational Behavior and Human Decision Processes*, 154, 80-95.
10.1016/j.obhdp.2019.09.003
- Kacmar, K. M., Tillman, C. J., Harris, K. J. et Whitman, M. V. (2015). Perceptions of ethical leadership as a mediator of the relationship between abusive supervision and work behaviors. *Journal of Behavioral & Applied Management*, 16(1), 2-19.
- Kant, L., Skogstad, A., Torsheim, T. et Einarsen, S. (2013). Beware the angry leader: Trait anger and trait anxiety as predictors of petty tyranny. *The Leadership Quarterly*, 24, 106-124.
<http://dx.doi.org/10.1016/j.leaqua.2012.08.005>
- Kelemen, T. K., Matthews, S. H. et Breevaart, K. (2020). Leading day-to-day: A review of the daily causes and consequences of leadership behaviors. *The Leadership Quarterly*, 31. doi:
10.1016/j.leaqua.2019.101344
- Kernan, M. C., Racicot, B. M. et Fisher, A. M. (2016). Effects of abusive Supervision, psychological climate, and felt violation on work outcomes: A moderated mediated model. *Journal of Leadership & Organizational Studies*, 23(3), 309-321. doi : 10.1177/1548051815627358
- Kiewitz, C., Restubog, S. L. D., Zagenczyk, T. J., Scott, K. D., Garcia, P. R. J. M. et Tang, R. L. (2012). Sins of the parents: Self-control as a buffer between supervisors' previous experience of family undermining and subordinates' perceptions of abusive supervision. *The Leadership Quarterly*, 23, 869–882. <https://doi.org/10.1016/j.leaqua.2012.05.005>
- Kim, K. Y., Atwater, L., Latheef, Z. et Zheng, D. (2019). Three motives for abusive supervision: The mitigating effect of subordinates attributed motives on abusive supervision's negative outcomes. *Journal of Leadership & Organizational Studies*, 26(4), 476-494. doi :
10.1177/1548051818781816
- Kline, T. J. B. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage Publications, Inc.
- Kluemper, D. H., Mossholder, K. W., Ispas, D., Bing, M. N., Iliescu, D. et Ilie, A. (2019). When core self-evaluations influence employees' deviant reactions to abusive supervision: The moderating role of cognitive ability. *Journal of Business Ethics*, 159(2), 435-453. doi : 10.1007/s10551-018-3800-y
- Krasikova, D. V., Green, S. G. et LeBreton, J. M. (2013). Destructive leadership: A theoretical review, integration, and future research agenda. *Journal of Management*, 39(5), 1308-1338. doi :
10.1177/0149206312471388
- Labrague, L. J., Loriga, J., Nwafor, C. E., van Bogaert, P. et Cummings, G. G. (2020). Development and psychometric testing of the toxic leadership behaviors of nurse managers (ToxBH-NM) scale. *Journal of Nursing Management*, 28(4), 840-850. doi : 10.1111/jonm.13008

- Lam, C. K., Walter, F. et Huang, X. (2017). Supervisors' emotional exhaustion and abusive supervision: The moderating roles of perceived subordinate performance and supervisor self-monitoring. *Journal of Organizational Behavior*, 38(8), 1151-1166. doi : 10.1002/job.2193
- Lang, P., Kozak, M., Miller, G., Levin, D. et McLean, A. (1980). Emotional imagery: Conceptual structure and pattern of somato-visceral response. *Psychophysiology*, 17, 179 –192. doi:10.1111/j.1469-8986.1980.tb00133.x
- Larsson, G., Brandebo, M. F. et Nilsson, S. (2012). Destrudo-L: Development of a short scale designed to measure destructive leadership behaviours. *Leadership & Organization Development Journal*, 33(4), 383-400.
- Lavoie-Tremblay, M., Fernet, C., Lavigne, G. L. et Austin, S. (2016). Transformational and abusive leadership practices: Impacts on novice nurses, quality of care and intention to leave. *Journal of Advanced Nursing*, 72(3), 582-592. doi :10.1111/jan.12860
- Lazarus, R. S. et Folkman, S. (1984). *Stress, appraisal, and coping*. Springer.
- Lee, J., Wang, G. et Piccolo, R. F. (2018). Jekyll and hyde leadership: A multilevel, multisample examination of charisma and abuse on follower and team outcomes. *Journal of Leadership & Organizational Studies*, 25(4), 399-415. doi : 10.1177/1548051818757692
- Liang, L. H., Lian, H., Brown, D. J., Ferris, D. L., Hanig, S. et Keeping, L. M. (2016). Why are abusive supervisors abusive? A dual-system self-control model. *Academy of Management Journal*, 59(4). doi : 1385-1406. 10.5465/amj.2014.0651
- Liao, Z., Yam, K. C., Johnson, R. E., Liu, W. et Song, Z. (2018). Cleansing my abuse: A reparative response model of perpetrating abusive supervisor behavior. *Journal of Applied Psychology*, 103(9), 1039-1056. <http://dx.doi.org/10.1037/apl0000319>
- Liao, Z., Lee, H. W., Johnson, R. E., Song, Z. et Liu, Y. (2020). Seeing from a short-term perspective: When and why daily abusive supervisor behavior yields functional and dysfunctional consequences. *Journal of Applied Psychology*. doi : 10.1037/apl0000508
- Lin, S.-H., Ma, J. et Johnson, R. E. (2016). When ethical leader behavior breaks bad: how ethical leader behavior can turn abusive via ego depletion and moral licensing. *Journal of Applied Psychology*, 101(6), 815-830. doi : 10.1037/apl0000098
- Lipman-Blumen, J. (2010). Toxic leadership: A conceptual framework. Dans F. Bournois, J. Duval-Hamel, S. Roussillon et J. Scaringella (dir.), *Handbook of top management teams* (p. 214-220). Londres, Royaume-Uni : Palgrave Macmillan.
- Low, Y. M., Sambasivan, M. et Ho, J. A. (2019). Impact of abusive supervision on counterproductive work behaviors of nurses. *Asia Pacific Journal of Human Resources*. doi : 10.1111/1744-7941.12234

- Lukacik, E.-R. et Bourdage, J. S. (2019). Exploring the influence of abusive and ethical leadership on supervisor and coworker-targeted impression management. *Journal of Business and Psychology*, 34(6), 771-789. 10.1007/s10869-018-9593-2
- Ma, H., Karri, R. et Chittipeddi, K. (2004). The paradox of managerial tyranny. *Business Horizons*, 47(4), 33-40.
- Mackey, J. D., Frieder, R. E., Brees, J. R. et Martinko, M. J. (2017). Abusive supervision: A meta-analysis and empirical review. *Journal of Management*, 43. doi:10.1177/0149206315573997
- Martinko, M. J., Harvey, P., Brees, J. R. et Mackey, J. (2013). A review of abusive supervision research. *Journal of Organizational Behavior*, 34(S1), S120-S137. doi:10.1002/job.1888
- Mead, N. L., Baumeister, R. F., Stuppy, A. et Vohs, K. D. (2018). Power increases the socially toxic component of narcissism among individuals with high baseline testosterone. *Journal of Experimental Psychology: General*, 147(4), 591-596.
- Mintzberg, Henry. (1979). *The Structuring of Organizations*. Englewood Cliffs, New Jersey : Prentice-Hall. Repéré à <https://www.nrc.gov/docs/ML0907/ML090710600.pdf>
- Mitchell, M. S. et Ambrose, M. L. (2007). Abusive supervision and workplace deviance and the moderating effects of negative reciprocity beliefs. *Journal of Applied Psychology*, 92(4), 1159-1168.
- Mitchell, M. S., Vogel, R. M. et Folger, R. (2015). Third parties' reactions to the abusive supervision of coworkers. *Journal of Applied Psychology*, 100(4), 1040-1055.
<http://dx.doi.org/10.1037/apl0000002>
- Nielsen, M. B., Skogstad, A., Matthiesen, S. B. et Einarsen, S. (2016). The importance of a multidimensional and temporal design in research on leadership and workplace safety. *Leadership Quarterly*, 27(1), 142-155. doi : 10.1016/j.leaqua.2015.08.003
- Neuman, J. H. et Baron, R. A. (1998). Workplace violence and workplace aggression: Evidence concerning specific forms, potential causes, and preferred targets. *Journal of Management*, 24, 391– 419. doi:10.1177/014920639802400305
- Ogunfowora, B., Weinhardt, J. M. et Hwang, C. C. (2019). Abusive supervision differentiation and employee outcomes: The roles of envy, resentment, and insecure group attachment. *Journal of Management*. doi : 10.1177/0149206319862024
- Ong, C. W., Roberts, R., Arthur, C. A., Woodman, T. et Akehurst, S. (2016). The leader ship is sinking: A temporal investigation of narcissistic leadership. *Journal of Personality*, 84(2), 237-247.
doi:10.1111/jopy.12155
- Pearce, C. et Sims, H. P. (2002). Vertical versus shared leadership as predictors of the effectiveness of change management teams: An examination of aversive, directive, transactional, transformational,

- and empowering leader behaviors. *Group Dynamics: Theory, Research, and Practice*, 6(2), 172-197. doi: 10.1037//1089-2699.6.2.172
- Peng, A. C., Schaubroeck, J. M. et Li, Y. (2014). Social exchange implications of own and coworkers' experiences of supervisory abusive. *Academy of Management Journal*, 57(5), 1385-1405. <http://dx.doi.org/10.5465/amj.2012.0080>
- Peng, A. C., Schaubroeck, J. M., Chong, S. et Li, Y. (2019). Discrete emotions linking abusive supervision to employee intention and behavior. *Personnel Psychology*, 72(3), 393-419. doi : 10.1111/peps.12310
- Pett, M. A., Lackey, N. R. et Sullivan, J. J. (2003) *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. SAGE Publications, Inc. <http://dx.doi.org/10.4135/9781412984898>
- Poirel, E. et Yvon, F. (2012). Le leadership distribué : l'épreuve des réformes éducatives. *Éducation et francophonie*, 40(1), 94-118. <https://doi.org/10.7202/1010148ar>
- Priesemuth, M. et Bigelow, B. (2020). It hurts me too! (or not?): Exploring the negative implications for abusive bosses. *Journal of Applied Psychology*, 105(4), 410-421. doi : 10.1037/apl0000447
- Priesemuth, M. et Schminke, M. (2019). Helping thy neighbor? Prosocial reactions to observed abusive supervision in the workplace. *Journal of Management*, 45(3), 1225-1251. doi : 10.1177/0149206317702219
- Pundt, A. et Schwarzbeck, K. (2018). Abusive supervision from an integrated self-control perspective. *Applied Psychology : An International Review*, 67(3), 473-497. <https://psycnet.apa.org/doi/10.1111/apps.12125>
- Richard, O. C., Boncoeur, O. D., Chen, H. et Ford, D. L. (2020). Supervisor abuse effects on subordinate turnover intentions and subsequent interpersonal aggression: The role of power-distance orientation and perceived human resource support climate. *Journal of Business Ethics*, 164(3), 549-563. doi : 10.1007/s10551-018-4019-7
- Robinson, M. D. et Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, 27, 1520 –1532. doi:10.1177/01461672012711012
- Rousseau, V. et Aube, C. (2018). When leaders stifle innovation in work teams: The role of abusive supervision. *Journal of Business Ethics*, 151(3), 651-664. doi : 10.1007/s10551-016-3258-8
- Rus, D., van Knippenberg, D. et Wisse, B. (2010). Leader power and leader self-serving behavior: The role of effective leadership beliefs and performance information. *Journal of Experimental Social Psychology*, 46, 922-933.
- Sager, K. L. Wallace, E. A., Jarrett, B. et Richey, J. A. (2015). Exploring the role of communication

- competence in abusive supervision: A test of a model linking downward communicative adaptability to downward abusive communication. *SAGE Open*, 1-8. doi : 10.1177/2158244015604689
- Samreen, F. et Rashid, M. A. (2018). Abusive supervision and subordinates' retaliation: The mediating role of OCB. *Pacific Business Review International*, 11(5), 94-102.
- Sasso, T. et Gonzalez-Morales, G. (2018). Centering the target of mistreatment in our measures. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 11(1), 107-112. <https://doi.org/10.1017/iop.2017.92>
- Schaubroeck, J. M., Peng, A. C. et Hannah, S. T. (2016). The role of peer respect in linking abusive supervision to follower outcomes: Dual moderation of group potency. *Journal of Applied Psychology*, 101(2), 267-278. doi : 10.1037/apl0000050
- Schmid, E. A., Verdorfer, A. P. et Peus, C. V. (2018). Different shades – different effects? Consequences of different types of destructive leadership. *Frontiers in Psychology*, 9, 1-16. doi: 10.3389/fpsyg.2018.01289
- Schmid, E. A., Verdorfer, A. P. et Peus, C. V. (2019). Shedding light on leaders' self-interest: Theory and measurement of exploitative leadership. *Journal of Management*, 45(4), 1401-1433. doi:10.1177/0149206317707810
- Schmidt, A. A. (2008). *Development and validation of the toxic leadership scale* [thèse de doctorat, University of Maryland]. Digital Repository at the University of Maryland. <http://hdl.handle.net/1903/8176>
- Schyns, B. et Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, 24(1), 138-158. doi : 10.1016/j.leaqua.2012.09.001
- Shao, P., Li, A. et Mawritz, M. (2018). Self-protective reactions to peer abusive supervision: The moderating role of prevention focus and the mediating role of performance instrumentality. *Journal of Organizational Behavior*, 39(1), 12-25. doi : 10.1002/job.2206
- Shaw, J. B., Erickson, A. et Harvey, M. (2011). A method for measuring destructive leadership and identifying types of destructive leaders in organizations. *The Leadership Quarterly*, 22, 575–590. <https://doi.org/cqkvgb>
- Shum, C., Gatling, A. et Tu, M. (2020). When do abusive leaders experience guilt? *International Journal of Contemporary Hospitality Management*, 1-18. <http://dx.doi.org/10.1108/IJCHM-05-2019-0474>
- Tabachnick, B. G. et Fidell, L. S. (1996). *Using Multivariate Statistics* (3^e éd.). Harper Collins.
- Tepper, B. J. (2000). Consequences of abusive supervision. *Academy of Management Journal*, 43(2), 178-190. doi:10.5465/1556375

- Tepper, B. J. (2007). Abusive supervision in work organizations: Review, synthesis, and research agenda. *Journal of Management*, 33, 261-289. doi:10.1177/0149206307300812
- Tepper, B. J., Duffy, M. K. et Breaux-Soignet, D. M. (2012). Abusive supervision as political activity: Distinguishing impulsive and strategic expressions of downward hostility. Dans G. R. Ferris et D. C. Treadway (dir.), *Politics in organizations: Theory and research considerations* (p. 191-212). Routledge.
- Tepper, B. J., Duffy, M. K., Henle, C. A. et Lambert, L. S. (2006). Procedural injustice, victim precipitation, and abusive supervision. *Personnel Psychology*, 59, 101-123.
- Tepper, B. J., Mitchell, M. S., Haggard, D. L., Kwan, H. K. et Park, H.-M. (2015). On the exchange of hostility with supervisors: An examination of self-enhancing and self-defeating perspectives. *Personnel Psychology*, 68, 723-758. <https://doi.org/10.1111/peps.12094>
- Tepper, B. J., Simon, L. et Park, H. M. (2017). Abusive supervision. *The Annual Review of Organizational Psychology and Organizational Behavior*, 4, 123-152. doi: 10.1146/annurev-orgpsych-041015-062539
- Thomas, D. L. et Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, 59, 291-297. doi:10.1037/0022-3514.59.2.291
- Thoroughgood, C. N., Tate, B. W., Sawyer, K. B. et Jacobs, R. (2012). Bad to the bone: Empirically defining and measuring destructive leader behavior. *Journal of Leadership & Organizational Studies*, 19, 230-255. doi:10.1177/1548051811436327
- Trepanier, S.-G., Boudrias, V. et Peterson, C. (2019). Linking destructive forms of leadership to employee health. *Leadership & Organization Development Journal*, 40(7), 803-814. doi : 10.1108/LODJ-04-2019-0168
- Tzafrir, S. S., Gur, A. B.-A. et Blumen, O. (2015). Employee social environment (ESE) as a tool to decrease intention to leave. *Scandinavian Journal of Management*, 31(1), 136-146. doi : 10.1016/j.scaman.2014.08.004
- van Knippenberg, D. (2011). Embodying who we are: Leader group prototypicality and leadership effectiveness. *Leadership Quarterly*, 22, 1078-1091. doi : 10.1016/j.leaqua.2011.09.004
- Vogel, R., Homberg, F. et Gericke, A. (2016). Abusive supervision, public service motivation, and employee deviance: The moderating role of employment sector. *Evidence-Based Hrm: a Global Forum for Empirical Scholarship*, 4(3), 214-231. doi : 10.1108/EBHRM-08-2015-0034
- Vogel, R. M., Mitchell, M. S., Tepper, B. J., Restubog, S. L. D., Hu, C., Hua, W. et Huang, J.-C. (2015). A cross-cultural examination of subordinates' perceptions of and reactions to abusive supervision. *Journal of Organizational Behavior*, 36(5), 720-745. doi : 10.1002/job.1984
- Waldman, D. A., Wang, D., Hannah, S. T., Owens, B. P. et Balthazard, P. A. (2018). Psychological and

- neurological predictors of abusive supervision. *Personnel Psychology*, 71(3), 399-421. doi : 10.1111/peps.12262
- Walsh, B. M., Matthews, R. A., Toumbeva, T. H., Kabat-Farr, D., Philbrick, J. et Pavisic, I. (2019). Failing to Be Family-Supportive: Implications for Supervisors. *Journal of Management*, 45(7), 2952-2977. 10.1177/0149206318774621
- Wang, G., van Iddekinge, C. H., Zhang, L. et Bishoff, J. (2019). Meta-analytic and primary investigations of the role of followers in ratings of leadership behavior in organizations. *Journal of Applied Psychology*, 104(1), 70-106. doi : 10.1037/apl0000345
- Wee, E. X. M., Liao, H., Liu, D. et Liu, J. (2017). Moving from abuse to reconciliation: A power-dependence perspective on when and how a follower can break the spiral of abuse. *Academy of Management Journal*, 60(6), 2352-2380. doi : 10.5465/amj.2015.0866
- Woestman, D. S. et Wasonga, T. A. (2015). Destructive leadership behaviors and workplace attitudes in schools. *NASSP Bulletin*, 99(2), 147-163. doi : 10.1177/0192636515581922
- Wright, T. A. et Sweeney, D. A. (2016). The call for an increased role of replication, extension, and mixed-methods study designs in organizational research. *Journal of Organizational Behavior*, 37, 480-486. doi : 10.1002/job.2059
- Xu, S., Martinez, L. R., Van Hoof, H., Tews, M., Torres, L. et Farfan, K. (2018). The impact of abusive supervision and co-worker support on hospitality and tourism student employees' turnover intentions in Ecuador. *Current Issues in Tourism*, 21(7), 775-790. doi : 10.1080/13683500.2015.1076771
- Yam, K. C., Fehr, R., Keng-Highberger, F. T., Klotz, A. C. et Reynolds, S. J. (2016). Out of control: A self-control perspective on the link between surface acting and abusive supervision. *Journal of Applied Psychology*, 101(2), 292-301. doi : 10.1037/apl0000043
- Yergeau, E. et Poirier, M. (2013). *SPSS à l'UdeS*. <http://spss.espaceweb.usherbrooke.ca>
- Yeung, E. et Shen, W. (2019). Can pride be a vice and virtue at work? Associations between authentic and hubristic pride and leadership behaviors. *Journal of Organizational Behavior*, 40(6), 605-624. doi : 10.1002/job.2352
- Yu, L. et Duffy, M. K. (2020). The whiplash effect: The (moderating) role of attributed motives in emotional and behavioral reactions to abusive supervision. *Journal of Applied Psychology*. <http://dx.doi.org/10.1037/apl0000810>
- Zellars, K. L., Tepper, B. J. et Duffy, M. K. (2002). Abusive supervision and subordinates' organizational citizenship behavior. *Journal of Applied Psychology*, 87(6), 1068-1076.
- Zhang, Y. et Bednall, T. C. (2016). Antecedents of abusive supervision: A meta-analytic review. *Journal of Business Ethics*, 139(3), 455-471.

- Zhang, J. et Liu, J. (2018). Is abusive supervision an absolute devil? Literature review and research agenda. *Asia Pacific Journal of Management*, 35, 719-744. <https://doi.org/10.1007/s10490-017-9551-y>
- Zhang, Y., Liu, X. et Chen, W. (2019). Fight and flight: A contingency model of third parties' approach-avoidance reactions to peer abusive supervision. *Journal of Business and Psychology*. <https://doi.org/10.1007/s10869-019-09650-x>
- Zuroff, D. C., Fournier, M. A., Patall, E. A. et Leybman, M. J. (2010). Steps toward an evolutionary personality psychology: Individual differences in the social rank domain. *Canadian Psychology*, 51(1), 58-66.

Annexe A

Explications des choix d'exclusion des construits

Construits exclus	Explication(s)
<i>Misuse of power</i>	Le construit est défini comme suit : « <i>willingness to use one's power to improve one's outcomes at the expense of others</i> » (Mead, Baumeister, Stuppy et Vohs, 2018, p.593). Même s'il est très similaire à la définition <i>petty tyrant</i> d'Ashforth (1994), il comporte des comportements qui ne sont pas nécessairement abusifs envers autrui (ex : accepter un pot-de-vin).
Leadership sombre (<i>dark leadership / dark triad</i>)	Ces échelles s'inspirent des manifestations sous-cliniques de certains troubles de la personnalité : narcissisme, psychopathie, machiavélisme. Les questionnaires mesurent des traits de personnalité (ex. : <i>I tend to want others to admire me</i> pour le narcissisme; <i>I tend to lack remorse</i> pour la psychopathie; et <i>I have used flattery to get my way</i> pour le machiavélisme; Jonason et Webster, 2010) et non des comportements de leadership où un leader tente d'influencer son employé (ex. : critique sévèrement ses employés qui font des erreurs; Cacciatore, 2015). Les outils mesurant ces manifestations sous-cliniques de troubles de la personnalité pourraient faire l'objet d'un essai en soi; vingt-deux outils ont été recensés dans le cadre de la préparation du présent essai.
Leadership autoritaire (<i>authoritarian leadership</i>)	Un leader autoritaire peut, par exemple, demander à ses subordonnés d'obéir à ses instructions, d'effectuer les tâches comme il le demande sans quoi ils seront sévèrement punis (Cheng, Chou, Wu, Huang et Farh, 2004). Toutefois, un leader peut être autoritaire sans utiliser des moyens abusifs, et c'est pourquoi les types de leadership tyrannique et autoritaire sont traités comme étant des construits différents dans la littérature (Gabler et al., 2014). Les concepts sont toutefois liés : le leadership autoritaire peut être un antécédant à la supervision abusive (Gabler et al., 2014).
Leadership aversif	Le leadership aversif comporte le recours aux réprimandes et à l'intimidation (Pearce et Sims, 2002). L'outil utilisé dans Pearce et Sims (2002) mesure de façon trop limitée les manifestations que peuvent prendre le leadership tyrannique. À titre d'exemple, un leader aversif pourrait se montrer intimidant sans nécessairement dévaloriser ses employés ou utiliser son autorité à des fins personnelles (alors que des items mesurant ces dimensions sont inclus, par exemple, dans le questionnaire de Cacciatore, 2015 mesurant le leadership tyrannique). Un leader aversif pourrait aussi réprimander ses employés lorsque leur performance laisse à désirer (Pearce et Sims, 2002), alors qu'un leader abusif serait plutôt susceptible de donner des réprimandes même si le travail de l'employé est bien fait ou à l'égard de choses qui sont hors du contrôle de l'employé (Cacciatore, 2015). Il y a donc une certaine notion d'imprévisibilité qui fait implicitement partie d'items mesurant le leadership abusif, que ne captent pas les items du leadership aversif.

Leadership dominant	Il s'agit d'un construit qui diffère de la supervision abusive. Le construit représente une affirmation de soi et une propension au leadership (ex : faire valoir ses idées) plutôt qu'un leadership abusif (Zuroff, Fournier, Patall et Leybman, 2010).
<i>Leader derailment</i>	Ce construit dépasse ce qui est compris dans la définition du leadership tyrannique. Le leadership « déraillé » inclut des comportements anti-organisationnel (ex : fraude) qui ne constituent pas du leadership en soi (Einarsen et al., 2007). Ces leaders « déraillent » de ce qui est attendu d'eux. Ceci peut aussi prendre plusieurs formes comme plafonner prématurément dans son cheminement de carrière, être rétrogradé ou licencié (Cullen, Gentry et Yammarino, 2015), ce qui n'est pas compris dans la définition du leadership tyrannique.
<i>Leader self-interest</i>	Ce construit mesure la propension d'un leader à mettre de l'avant ses propres intérêts (Rus, van Knippenberg et Wisse, 2010). Même si ce construit est souvent associé à une manifestation qui nuit aux autres (ex : <i>I have shifted the blame for a mistake of mine onto one of my subordinates</i>), ce n'est pas nécessairement abusif envers les subordonnés de façon directe (ex : <i>I negotiated a bonus for myself that was substantially higher than the bonus my subordinates received</i>). Ceci est un exemple où ce peut être injuste, mais pas nécessairement de l'abus directe envers les employés. C'est donc un construit qui diverge du leadership tyrannique.
<i>Self-centered leadership</i>	Les leaders égocentriques (<i>self-centered leaders</i>), tel que le construit est défini par Hawass (2019), ne sont habituellement pas abusifs ou physiquement hostiles envers les employés (même si un leader égocentrique peut aussi être abusif, le construit tel que défini par l'auteur inclus les leaders égocentriques qui ne sont pas abusifs). Cela va à l'encontre de la définition de la supervision abusive.
Leadership toxique	La définition du construit est plus large que celle du leadership tyrannique. Elle inclut par exemple des activités qui nuisent à l'organisation, comme la fraude, et qui ne constituent pas nécessairement des actes de leadership (Lipman-Blumen, 2010). En ce sens, l'outil de Labrague, Lorica, Nwafor, van Bogaert et Cummings (2020) mesurant le leadership toxique n'a pas été retenu.
Leadership destructeur	Les échelles de leadership destructeur ne comportant pas de sous-échelle mesurant spécifiquement le leadership tyrannique ou la supervision abusive ou n'étant pas libellé comme visant les comportements dirigés envers les subordonnés n'étaient pas retenues dans la présente étude, étant donné que le leadership destructeur peut inclure des comportements dirigés envers l'organisation (Einarsen et al., 2007; Krasikova et al., 2013), ce qui ne fait pas partie de la définition du leadership tyrannique et de la supervision abusive du présent article. Ainsi, l'outil de Shaw, Erickson et Harvey (2011) a été exclus.
<i>Unethical leadership</i>	La définition du construit est plus large que celle du leadership tyrannique. Elle comporte des comportements et des décisions faits par des leaders qui sont illégaux et qui violent les principes moraux (Brown et Mitchell, 2010). Ceux-ci ne sont pas nécessairement des comportements de leadership abusif pour les employés (Krasikova et al., 2013).

<i>Downward abusive communication</i>	L'échelle de Tepper a été adaptée pour mesurer des comportements abusifs spécifiquement liés à la communication (Sager et al., 2015). Ceci dit, la supervision abusive peut s'exprimer autrement que par la communication (ex. : <i>invades my privacy</i>). Donc, cette échelle s'avère incomplète pour couvrir le phénomène.
<i>Workplace violence</i>	Le <i>Workplace violence</i> inclut les manifestations d'agression physique (Newman et Baron, 1998), ne faisant pas partie de la définition de la supervision abusive de Tepper (2000).

Annexe B

Précisions additionnelles concernant les analyses statistiques

Vérifications à l'égard des études exclues des analyses statistiques

Tel qu'expliqué dans la section des résultats, 242 études ont été retenues pour les analyses statistiques de l'objectif 1. Parmi celles-ci, vingt études ($k = 20$) ne précisait pas quel type d'échelle les chercheurs ont utilisé. C'est pourquoi il est indiqué dans le Tableau 3 que les corrélations présentées sont calculées à partir de 222 à 242 échantillons. Afin de vérifier si ces vingt études ayant des données manquantes diffèrent de celles n'ayant pas de données manquantes, une variable dichotomique a été créée (groupe 1 = donnée non manquante, type d'échelle indiqué; groupe 2 = donnée manquante, type d'échelle non indiqué) et des corrélations ont été calculées entre cette variable dichotomique et les caractéristiques mesurées par la présente recension des études (nombre d'items utilisé, nombre de points d'ancrage, indice de fidélité, nombre de répondants et région d'appartenance (Orient c. Occident) de l'échantillon). Les coefficients de corrélation (tels que rapportés dans le Tableau B1) étaient tous non significatifs ($-0,04 < r < 0,07$, $p > 0,05$), indiquant que les études n'ayant pas précisé le type d'échelle ne diffèrent pas des études l'ayant précisé.

Tableau B1

Coefficient de corrélation (r) entre la variable dichotomique regroupant les études ayant indiqué le type d'échelle utilisé en un groupe et ceux ne l'ayant pas indiqué en un autre groupe

	Nombre d'items	Nombre d'ancrages	Indice de fidélité (α)	Nombre de répondants	Région d'appartenance de l'échantillon
Variable dichotomique	-0,03	-0,04	-0,02	-0,04	0,07

Note. $k = 196-242$. La variable dichotomique a été codée par 1 pour ceux ayant indiqué le type d'échelle et par 2 pour ceux n'ayant pas indiqué le type d'échelle. Aucun des coefficients de corrélation n'était significatif à $p < 0,05$. $\alpha =$ alpha de Cronbach

De surcroît, parmi les 222 études dans lesquelles le type d'échelle était spécifié, vingt-cinq d'entre elles ($k = 25$) ont eu recours à un autre nombre d'items que cinq ou quinze. Ainsi, en traitant la variable du nombre d'items de façon catégorielle (groupe 1 = 5 items; groupe 2 = 15 items) plutôt que de façon continue pour effectuer l'ANOVA factorielle, les vingt-cinq études ont été exclues des analyses. La même procédure que celle susmentionnée fut effectuée : une variable dichotomique fut créée (groupe 1 = étude ayant utilisé 5 ou 15 items; groupe 2 = étude ayant utilisé un autre nombre d'items) et des analyses corrélationnelles furent réalisées entre cette variable dichotomique et les caractéristiques des études recensées (type d'échelle, nombre de points d'ancrage, indice de fidélité, nombre de répondants et région d'appartenance (Orient c. Occident) de l'échantillon). Les coefficients de corrélation (tels que rapportés

dans le Tableau B2) étaient tous non significatifs ($-0,11 < r < 0,05$, $p > 0,05$), indiquant que les études ayant utilisé cinq ou quinze items ne diffèrent pas des études ayant eu recours à un autre nombre d'items.

Tableau B2

Coefficient de corrélation (r) entre la variable dichotomique regroupant les études ayant utilisé cinq ou quinze items en un groupe et celles ayant utilisé un autre nombre d'items en un autre groupe

	Type d'échelle	Nombre d'ancrages	Indice de fidélité (α)	Nombre de répondants	Région d'appartenance de l'échantillon
Variable dichotomique	0,04	-0,04	-0,11	0,03	0,05

Note. $k = 196-242$. La variable dichotomique a été codée par 1 pour les études ayant eu recours à cinq ou quinze items et par 2 pour celles ayant utilisé un autre nombre d'items. Aucun des coefficients de corrélation n'était significatif à $p < 0,05$. $\alpha =$ alpha de Cronbach.

Statistiques descriptives

En plus des informations statistiques rapportées dans le texte du présent essai, le Tableau B3 et le Tableau B4 présentent quelques résultats descriptifs additionnels en lien avec les ANOVA.

Tableau B3

Statistiques descriptives de l'ANOVA réalisée afin de répondre à l'objectif 1

Nombre d'items	Type d'échelle	Moyenne de supervision abusive	Écart-type	Taille d'échantillon
5	Fréquence	1,93	0,89	32
	Accord	2,01	0,54	40
	Total	1,97	0,72	72
15	Fréquence	1,68	0,48	75
	Accord	2,30	0,83	50
	Total	1,93	0,71	125
Total	Fréquence	1,76	0,64	107
	Accord	2,17	0,73	90
	Total	1,95	0,71	197

Tableau B4

Statistiques descriptives de l'ANOVA réalisée afin de répondre à l'objectif 6

Type d'échelle	Région	Moyenne de la supervision abusive	Écart-type	Taille d'échantillon
Fréquence	Occident	1,55	0,43	54
	Orient	2,01	0,77	48
	Total	1,76	0,65	102
Accord	Occident	1,62	0,36	23
	Orient	2,38	0,74	62

	Total	2,18	0,74	85
Total	Occident	1,57	0,41	77
	Orient	2,22	0,77	110
	Total	1,95	0,72	187

Vérification des postulats

Une ANOVA à devis factoriel a pu être effectuée étant donné que les postulats étaient rencontrés (Yergeau et Poirier, 2013, repéré à <http://spss.espaceweb.usherbrooke.ca/pages/statistiques-inferentielles/analyse-de-variance.php>). Premièrement, la **normalité** est respectée; les indices d'asymétrie et de Kurtosis de la courbe des moyennes de supervision abusive se chiffrent respectivement à 1,48 et à 1,59 (c'est-à-dire entre -2 et 2; Yergeau et Poirier, 2013, repéré à <http://spss.espaceweb.usherbrooke.ca/pages/stat-descriptives/procedure-explorer.php>). Deuxièmement, les échantillons sont **indépendants** les uns des autres; une attention particulière a été portée à cet égard lors de la sélection des études et de l'extraction des données, tel que mentionné antérieurement dans le texte du présent essai. Troisièmement, l'**homogénéité** est respectée (Tabachnick et Fidell, 1996) : la comparaison de la taille des groupes est en deçà du ratio limite de 4 pour 1 (aucun groupe n'étant quatre fois plus grand qu'un autre); et la comparaison de la variance des groupes est aussi en deçà du ratio limite de 10 pour 1 (la variance d'un groupe n'étant jamais dix fois plus grande que la variance d'un autre groupe). Effectivement, pour l'ANOVA de l'objectif 1, le ratio obtenu pour la taille des groupes est de 1,74 pour le nombre d'items et de 1,19 pour le type d'échelle. Le ratio obtenu pour la variance des groupes est de 1,03 pour le nombre d'items et de 1,30 pour le type d'échelle. Pour l'ANOVA de l'objectif 6, les ratios de la taille des groupes sont de 1,20 pour le type d'échelle et de 1,43 pour l'appartenance à une région; les ratios de la variance des groupes sont de 1,30 pour le type d'échelle et de 3,53 pour l'appartenance à une région. Somme toute, les trois postulats sont rencontrés.

Analyse de puissance

Pour de telles analyses composées de quatre groupes (ANOVA de l'objectif 1 : deux groupes pour le nombre d'items et deux groupes pour le type d'échelle; ANOVA de l'objectif 6 : deux groupes pour le type d'échelle et deux groupes pour l'appartenance à une région), un minimum de quarante-cinq échantillons par groupe donnerait 80 % de chance de détecter statistiquement un effet moyen (lorsqu'il y a un effet réel), tout en utilisant un seuil alpha de 0,05 (Cohen, 1992). Aucun des groupes n'étant en deçà de 45 échantillons, la puissance permettant de détecter un effet significatif moyen n'est pas problématique.