Development and validity of the Ethical Leadership Questionnaire

Lyse Langlois
*Industrial Relations, Laval University,
Quebec City, Canada*
Claire Lapointe and Pierre Valois
*Educational Foundations and Practice, Laval University,
Quebec City, Canada, and*
Astrid de Leeuw
University of Luxemburg, Luxembourg City, Luxembourg

Abstract

Purpose – This study had five objectives: explain the initial steps that led to the construction of the Ethical Leadership Questionnaire (ELQ); analyze the items and verify the ELQ reliability using item response theory (IRT); examine its factorial structure with a confirmatory factor analysis (CFA) and an exploratory structural equation modeling (ESEM) approach; test the item bias of the ELQ; assess the relation between the ELQ dimensions and ethical sensitivity. The paper aims to discuss these issues.

Design/methodology/approach – Study 1 and Study 2 involved 200 and 668 respondents, respectively. Step 1 consisted in IRT; Step 2 in CFA and ESEM analysis; Step 3 in invariance of the ELQ items across gender, and Step 4 in structural equation modeling.

Findings – Results indicated the presence of the three types of ethic in the resolution of moral dilemmas, validating Starratt's model. The factor structure was gender invariant. Ethic of critique was significantly related to ethical sensitivity.

Research limitations/implications – More replications will be needed to fully support the ELQ's validity. Given that the instrument may be used in diverse cultural contexts, invariance across cultures would be warranted.

Practical implications – As educational organizations become aware of the crucial need for more ethical leaders, they will need to pay particular attention to the ethic of critique as it appears to play a significant role in the development of ethical sensitivity.

Social implications – Results presented in this paper answer a vital need for more ethical skills in educational leadership.

Originality/value – The ELQ provides a validated measure of Starratt's conceptual framework and highlights the key role played by ethical sensitivity and the ethic of critique.

Keywords Gender, Measurement, Leadership development, Evaluation, Educational administration, Skills, Ethical leadership

Paper type Research paper

Introduction

Recently, corruption in business and politics has contributed to bringing ethics to the forefront of public awareness by highlighting the effects of unethical behavior on organizations. Consequently, an increasing number of researchers have become interested in understanding the ethical dimensions of leadership. Our own research program aims at understanding the role different ethical dimensions play in the exercise of ethical leadership. This research program has led to a better understanding of ethical leadership in education (Langlois, 1997, 2004), to the identification of ethical sensitivity as a component of ethical conduct

(Langlois and Lapointe, 2009), to the development and experimentation of a training program for ethical leadership (Langlois and Lapointe, 2010), as well as to the conception of Ethical Leadership Questionnaire (ELQ), a questionnaire measuring ethical leadership as conceptualized by Starratt (1991). This last step of our research program aimed at contributing to what Starratt (2012) calls the cultivation of ethical schools. In this paper, we present the process which led to the development and validation of the ELQ.

Problem statement
In the 1970s and 1980s, interest in the moral dimensions of educational administration grew out of work by Bates (1982), Foster (1986), Greenfield (1987), Halpin (1970), and Hodgkinson (1978) who insisted on the need for a repositioning of leadership on axiological foundations. Starratt, in his seminal paper published in 1991, took a major step in this direction by incorporating moral reasoning and critical theory constructs into his model of ethical leadership. A few years later, Beck (1994) adopted a position based on an ethic of care while Brunner (1998) found that two ethical components could be identified in the way educational leaders work: the ethic of justice and the ethic of care. These initial studies were part of a growing body of literature in educational administration (e.g. Beck, 1994; Begley and Johansson, 2003; Cranston *et al.*, 2005;Crowson, 1989; Enomoto and Kramer, 2007; Kirby *et al.*, 1992; Marshall *et al.*, 1993; Maxcy, 2002; Stefkovich and Shapiro, 1995[1]; Strike *et al.*, 1998), which provided a framework for studying the characteristics of an ethics-oriented practice of educational leadership.

More recently, while concern for ethical leadership in the field of business administration emerged in reaction to scandals, corruption and conflicts of interest (Colvin, 2003; Kalshoven, 2010; Kalshoven *et al.*, 2011; Mehta, 2003), interest in ethics in educational administration became more pronounced given the need for vigilance in the interests of greater social justice, a value that is highly cherished in the world of education. This vigilance was in response to cuts in education budgets and their effects on youth as well as to excessive monitoring leading to an ever-increasing technologization of management practices (Bates, 1982; Stiggins, 2004). These two phenomena raised fears that the rationalization process underway was such that no one, either within or outside the school system, would be able to grasp its implications.

Until now, ethical leadership in education has been studied mostly through qualitative research, an approach which has provided rich information on this complex phenomenon. However, such qualitative inquiry is limited in its capacity to identify the key variables in the actualization of ethical conduct which justifies the need for quantitative studies of ethical leadership.

Measuring ethical leadership
Brown *et al.* (2005) developed the first instrument to measure ethical leadership in business administration, the Ethical Leadership Scale (ELS). In their paper, Brown and his colleagues mention that the ELS constitutes an initial step in the measurement of ethical leadership and that certain limits are still present. We agree with these authors for the following reasons. First, ELS items were tested mainly with participants from multi-unit financial services firms and MBA and psychology students. This does not take into consideration the fact that using psychology students as participants may create a bias given their lack of experience

in management. Second, in Brown *et al.*'s (2005) instrument, moral reasoning is associated with a transformational framework where ethical leadership is defined as the:

> […] demonstration of normative appropriate conduct through personal actions and interpersonal relationships, and the promotion of such conduct to followers through two-way communication, reinforcement and decision-making (p. 120).

However, recent advances in organizational and professional ethics stress the importance of exercising moral judgment, an aspect which is absent from the ELS.

More recently, Kalshoven *et al.* (2011) developed a multi-dimensional Ethical Leadership at Work (ELW) questionnaire based on seven behaviors related to ethical leadership in business administration (fairness, integrity, ethical guidance, people orientation, power sharing, role clarification, and concern for sustainability). The ELW is in line with Brown *et al.*'s (2005) model and aims at assessing the conduct of managers based on these seven elements. Kalshoven *et al.* (2011) affirm that their questionnaire can help to anticipate and gain a better understanding of the outcomes of a decision made when exercising ethical leadership. However, they acknowledge that the process involved in validating their tool is not yet complete.

In education, Edmonson *et al.* (2003) created an instrument to measure ethical leadership based on a normative perspective of ethics aiming at distinguishing "right" from "wrong" behaviors among educational administrators. Their instrument is framed on a Kantian deontological vision which aims at reinforcing good conduct and punishing bad conduct. Such a normative perspective of ethics fundamentally differs from our own in that we define ethics as a reflexive capacity that allows leaders autonomy in their ethical judgment. In fact, although a normative perspective is required to control professional conduct, our own results indicate that a capacity for ethical reflection is essential when confronted with the grey areas in ethical leadership (Langlois and Lapointe, 2010). Eyal *et al.* (2011) later introduced the Ethical Perspective Instrument (EPI), an exploratory tool which examines the ethical considerations involved in school leadership decision making. This questionnaire uses short hypothetical scenarios of core ethical dilemmas (EDs) school principals may experience. According to Eyal *et al.*, while their results provided valuable insight into the complex value system which underlies educational leadership today, they also found a need to supplement the EPI with research anchored in real-world contexts. This observation is supported by Lavorata (2005), Liedtka (1992), and Treviño and Weaver (1998) who underlined the limitations of scenarios or vignettes in the study of ethical leadership and conduct.

According to Yukl *et al.* (2013), given the absence of a clear definition of ethical leadership, confusion exists in the literature where it is sometimes associated with transformational leadership, sometimes with servant leadership. In our own research, ethical leadership is defined as a social practice by which professional judgment is autonomously exercised. It constitutes a resource rooted in three ethical dimensions – critique, care, and justice– as well as a powerful capacity to act in a responsible and acceptable manner.

Prompted by these observations, we developed and validated an instrument, the ELQ, which is framed in a descriptive ethical process inspired by Starratt's (1991) tridimensional model. The ELQ items aim at triggering a reflection on one's own professional conduct when faced with real EDs at work.

Starratt's tridimensional model
In his early work, Starratt (1991) first defined ethical leadership as based on three interdependent dimensions of ethics (justice, critique, and care) and later suggested that ethical leadership occurs through a meaningful process of learning (Starratt, 2004, 2012). This vision is framed in an ethics of virtues as envisioned by Aristotle where ethical leadership is seen as being part of a process of moral character development and increased awareness of ethical issues. The ELQ provides a measurement for the presence of each of the three ethics as defined in the following paragraphs.

Ethic of critique. The tenets of this ethical dimension examine injustice present in social relations or created by laws, within the structure of an organization or through the use of language. In short, this ethic reveals perspectives that could benefit one person or one group to the detriment of another person or another group. When an injustice is discovered, those who adopt this ethical perspective tend to sensitize others in order to obtain a better balance in the distribution of social benefits. The ethic of critique can be linked to Kohlberg's (1972, 1981) post-conventional level wherein a person might disobey a rule that he or she considers inequitable. Writing about leaders who act according to this ethic, Starratt (1991) states that "their basic stance is ethical for they are dealing with questions of social justice and human dignity, although not with individual choices" (p. 189).

Ethic of justice. The ethic of justice is rooted in the practice within a community and the assumption that the protection of human dignity depends on the moral quality of social relations which, in the end, is a public and political matter. According to Starratt (1991), in this perspective:

> [...] a communal understanding of the requirements of justice and governance flows from both tradition and the present effort of the community to manage its affairs in the midst of competing claims of the common good and individual rights. That understanding is never complete; it will always be limited by inadequacy of tradition to respond to changing circumstances and by the impossibility of settling conflicting claims conclusively and completely. The choices, however, will always be made with sensitivity to the bonds that tie individuals to their communities (p. 193).

The goal of the ethic of justice advocated by Starratt (1991) is to provoke exchanges, engage in debate and demonstrate transparency in management and to look for solutions when understanding is challenged. Consultation is the favored strategy to promote a positive attitude toward and an understanding for this type of ethic. The intention of those who act in accordance with an ethic of justice is to aim for responsible autonomy based on some form of cooperation. This intention consists of promoting a just social order within the organization as a result of collaboration between all people involved.

Ethic of care. Starratt's (1991) description of the ethic of care is based on the work of Gilligan (1982) and Noddings (1984). The ethic of care relates to the requirement of interpersonal relations, not from a contractual or legal standpoint, but in terms of absolute respect. Starratt (1991) adds that:

> This ethic places the human persons-in-relationship as occupying a position for each other of absolute value; neither one can be used as a means to an end; each enjoys an intrinsic dignity and worth, and given the chance, will reveal genuinely loveable qualities. An ethic of caring requires fidelity to persons, a willingness to acknowledge their right to be who they are, an openness to encountering them in their authentic individuality, a loyalty to the relationship (p. 195).

The intention of those who support an ethic of care is to consider human relations as being of major importance in the proper functioning of organizations.

The construct of ethical sensitivity

Moral theory of ethical judgment in the workplace underlies the concept of ethical sensitivity. The special function of this theory is to provide individuals with conceptual guidance for choosing to act in situations where there are conflicting moral claims or EDs. According to Rest (1986), moral judgment is based on four dimensions: ethical sensitivity, ethical judgment, ethical motivation, and ethical character. Rest (1986) and Cooper (2006) agree that ethical sensitivity should be recognized as the first step in ethical behavior development whereas Tuana (2007) posits that ethical sensitivity is a key element in moral literacy development which involves at least three major components:

(1) the ability to determine whether or not a situation involves ethical issues;

(2) awareness of the moral intensity of the ethical situation; and

(3) the ability to identify the moral virtues or values underlying an ethical situation.

Clarkeburn (2002) has similarly argued for the primacy of ethical sensitivity writing that "without recognizing the ethical aspects of a situation, it is impossible to solve any moral/ethical problem, for without the initial recognition, no problem exists" (p. 439). Rest (1986) has labeled this skill "moral sensitivity" whereas Butterfield *et al.* (2000) refer to it as "moral awareness," Damasio (1999) as "moral conscience," and Cooper (2006) as the "perception of an ethical problem."

However, measurement studies of ethical leadership focus mainly on the last three dimensions of Rest's model and little on ethical sensitivity, as is the case with questionnaires proposed by Riggio *et al.* (2010), Walumbwa *et al.* (2008), Yukl *et al.* (2013), and Zheng *et al.* (2011), Clarkeburn's (2002) being the only exception. Yet, our own findings point to the presence of a clear link between ethical sensitivity and ethical leadership. According to our action research studies, developing ethical leadership requires the activation of ethical sensitivity in order to better exercise one's moral judgment (Langlois and Lapointe, 2010). Thus, based on our previous research, we hypothesized that the three dimensions of the ELQ will be significantly related to ethical sensitivity (Langlois and Lapointe, 2007, 2009, 2010).

Research objectives

The aim of this paper is thus to: explain the initial steps that led to the construction of ELQ, analyze the items and verify the reliability of the ELQ using item response theory (IRT), examine the factorial structure of the ELQ with a confirmatory factor analysis (CFA) and an exploratory structural equation modeling (ESEM) approach, test the item bias of the ELQ, and assess the relation between the ELQ dimensions and ethical sensitivity.

Method

Pilot study: development of the ELQ

The questionnaire is based on data gathered from 200 semi-structured interviews with educational leaders (55 percent male and 45 percent female) carried out between 1995 and 2005 (Langlois, 1997, 2004). Through a qualitative approach, we were able to understand the real-life ethical aspects of educational leadership and to better identify the nature of school leaders' ethical challenges. The concept of ED is the core unit we used to assess the demonstrated presence of ethical leadership. The following definition of an ED was presented to participants: a person is faced with an ED when he or she experiences a major conflict between a number of values that are equally

important to him or her. This conflict prevents him or her from making a decision. According to Hatcher (1998):

> In resolving a moral dilemma, we can do no better than to evaluate carefully the various elements of the interaction, to determine the predominant principles, and then to act accordingly. We will not always succeed, even when trying our best, nor will we always have the time to engage in extensive reflection before acting (p. 121).

Prior to the interviews, a letter was sent to the volunteers explaining what we meant by ED and asking the person to reflect on such a dilemma they had encountered and were able to resolve to their satisfaction, or not. At the beginning of each interview, we reminded the participant what we meant by an ED and asked them to carefully describe the real-life ED they had chosen as well as the steps taken by them to successfully resolve it, or the obstacles encountered for those who failed to do so. The interviews were transcribed in their entirety and sent to each participant to validate their content. Once verified, we proceeded with a thematic analysis (Miles and Huberman, 2003; Paillé and Muchielli, 2003) based on Starratt's (1991) conceptual framework.

The first stage of the analysis consisted of identifying the types of EDs educational administrators encounter in their work. The second stage was to identify moral behaviors or ethical conducts that enabled participants to resolve their dilemma and to associate their self-reported behavior or conduct with the ethical constructs of care, justice, and critique, as defined by Starratt. These first two steps were used to verify Starratt's (1991) tridimensional model and to design a typology of moral behaviors associated with each of the three ethics. Indeed, despite the fact that justice and care were well-known dimensions in applied ethics, at the time this research was conducted, there were no known typologies to identify moral actions associated with each of them, except for an initial proposition by Brown *et al.* (1988) which inspired our work.

Identifying ethical behaviors also allowed us to expand the ethic of critique, which was not as well documented in the literature as the other two dimensions. The typology thus obtained contained 13 moral behaviors related to the ethic of care, eight related to the ethic of justice and nine related to ethic of critique, resulting in 30 items in total. The following are examples of these items: "I try to make people aware that some situations disproportionately privilege certain people" (ethic of critique); "I establish trust in my relationships with others" (ethic of care); "I try to be fair" (ethic of justice). The typology was then applied to build up a corpus of moral behaviors that initiate a process of moral reflection. It must be noted that, for some of the items, it was difficult to clearly identify to which main dimension they belonged.

More specifically, the purpose of the second stage of analysis was to carefully identify moral actions undertaken by leaders who were satisfied that they had successfully resolved their ED. This stage allowed us to identify the similarities among steps taken by those who felt successful in resolving their dilemma and led to the modeling of an ethical decision-making process which was then integrated in the questionnaire through five components: identifying an ED, solving an ED, typifying decisions made when facing an ED, assessing the influence of organizational culture on the process, and identifying pressures felt while resolving an ED. For each component, response items were linked to one of the three ethics.

In the main study, as described in the next section, we then initiated the third phase of the development of the questionnaire, which was testing the validity of the ELQ. To attain this objective, four steps were taken. In the first step, we used IRT to verify the psychometric properties of the ELQ. More specifically, we examined each item and

option characteristic curves, item parameters (discriminating power and the item difficulty or the degree of leadership needed to endorse an item) and item informativeness. In the second step, we applied the traditional CFA approach and the more recent ESEM approach (Asparouhov and Muthén, 2009; Marsh *et al.*, 2010) to test the proposed three-factor structure of the ELQ. The aim of the third step was to determine whether gender bias was present or not. We examined differential item functioning (DIF) to distinguish group mean differences with respect to how men and women endorse leadership in the ELQ. Finally, the aim of the fourth step was to test the performance of the three dimensions of the ELQ in predicting ethical sensitivity.

Main study: psychometric analysis of the ELQ
Participants and procedure of data collection. Data were collected from a sample of 668 North American educational leaders, 50.3 percent male and 49.7 percent female. Almost half of the male ($n = 135$; 40.7 percent) and the female respondents ($n = 153$; 45.5 percent) were aged between 46 and 56. Out of the 668 participants, 381 were Canadian secondary school principals, 259 were Canadian community college administrators and 28 were American high school principals. The data were collected using two strategies: 413 participants completed a paper version of the ELQ and 255 participants responded to an online version of the ELQ.

In order to improve the validity of our results and respect the probability assumption that each of the four data analysis steps described above is more independent of the others, we decided to divide our sample into four random subgroups of 167 participants. Thus, each block of data analyses, corresponding respectively to the four objectives of the study, was conducted on a different subgroup of participants.

Measures
ELQ. The original version of the ELQ contained the 30 items formulated in the pilot study in such a way to measure the following constructs: the ethic of care (13 items), the ethic of justice (eight items) and the ethic of critique (nine items). For instance, the following items were used to measure these three dimensions: "When I reflect on how I act at work, my main concern is to maintain harmony in the organization" (ethic of care); "When I have to solve an ethical problem, I try to clarify the main obstacles to understanding the situation" (ethic of critique); "When I have to solve an ethical problem, I check out the legal implications and any regulations that might apply" (ethic of justice). As stated above, it was difficult to clearly identify to which main dimension some items were related. Respondents were asked to rate each statement on a six-point Likert scale with response options ranging from 1 (never) to 6 (always). The statistical analysis strategy used to test the psychometric properties of the ELQ is presented after the description of the measure "ethical sensitivity."

Ethical sensitivity. We measured this variable because we hypothesized that the three dimensions of the ELQ would be significant predictors of ethical sensitivity (Langlois and Lapointe, 2010). Three items were used to measure ethical sensitivity. The participants were asked to rate each statement on a six-point Likert scale with response options ranging from 1 (never) to 6 (always): "For me, a situation that produces inequality presents an ethical dilemma," "For me, a situation which involves power-tripping creates an ethical dilemma," and "For me, situations that are hurtful to people create ethical dilemmas." This measure had good internal consistency (Cronbach's $\alpha = 0.80$).

Statistical analysis strategy

Step 1: IRT. IRT refers to a set of mathematical models that describe, in probabilistic terms, the relationship between a person's response to an item of a scale, test or questionnaire, and his or her level of the latent variable, such as ethic of justice, being measured (Reeve and Fayers, 2005). The fundamental objective of IRT methods is thus to specify the relationships between item responses and the latent trait (θ) posited to be measured by the test or questionnaire. An item characteristic curve (ICC) graphically represents the manner in which the probability of a response varies with the level of the underlying trait. If the probability of endorsing an item or option increases as a function of, for instance, ethical leadership, the item is effective.

Several models of IRT can be applied to rating data but, given the use of Likert scales, the graded-response model (GRM; Samejima, 1969, 1997) was deemed most appropriate (Embretson and Reise, 2000; Poitras *et al.*, 2012). In order to compute the parameter estimates under the GRM, the EIRT program (Excel IET assistant; Valois *et al.*, 2011) was used in the current study. More specifically, it was used to estimate the ability of items and scales to detect differences between individuals across a latent trait. Ideally, the differences between two individuals should be as easily detected, irrespective of whether they are in the low or high range in terms of ethical leadership. The slope of the ICC (a_i; item discriminability) indicates the extent to which a change in item score corresponds to changes in the level of underlying latent trait (Baker, 2001). Hence, the larger the value of a_i, the steeper will be the slope.

After analyzing the degree of discrimination, we examined the b_{ij} parameters (one for each category response curve) and their standard errors. The b_{ij} parameters represent a given leadership category level as a function of the maximal probability of choosing each response option. For intermediate options, the threshold modal is the place on the latent trait (i.e. ethical leadership) scale, where the probability of choosing this option is maximal. For the first option, that is the place on the theta scale where there is a 50 percent chance to choose the first option, this probability increases to the left and decreases to the right. For the last option, that is the place on the theta scale where there is a 50 percent chance of choosing the last option, this probability decreasing to the left and increasing to the right. Good items have b_{ij} parameters that match their respective response categories and have a low error level (Poitras *et al.*, 2012). Finally, item informativeness (I_i) was also used in order to assess the accuracy of the ethical leadership level estimation. The higher the amount of information, the more accurately the latent trait level (ethical leadership) is estimated (Baker, 2001).

Based on Baker (2001), we decided to keep the items with an $a_i > 0.65$, which corresponds to a moderate differentiation. After having analyzed the degree of discrimination, the b_{ij} parameters (one for each category response curve) and their standard errors were considered. We kept items with b_{ij} parameters having an adequate location on the scale continuum and acceptable standard errors. Finally, item informativeness (I_i) was also used in order to assess the accuracy of the ethical leadership level estimation. As noted above, the higher the amount of information, the more accurately the latent trait level (e.g. ethic of critique, ethic of justice, and ethic of care) is estimated.

Step 2: CFA and ESEM analysis. In CFA studies, each item is allowed to load on one factor, even though many measurement instruments show small cross-loadings that are well motivated by either substantive theory or by the formulation of the measurements (Asparouhov and Muthén, 2009; Marsh *et al.*, 2009). There are some situations in which this strict requirement of zero cross-loadings does not fit the data

well, leading to distorted factors and overestimated factor correlations (Marsh *et al.*, 2009). To overcome this limitation, Asparouhov and Muthén (2009) proposed a new statistical method that integrates CFA and EFA: ESEM. This method allows for the use of rotation criteria within structural equation modeling (SEM) analysis. In this study, both CFA and ESEM methods were used to examine the factorial structure of the ELQ. They were performed using Mplus 6 software (Muthén and Muthén, 1998/2010). As Marsh *et al.* (2009) recommend, we began with a CFA analysis to verify the appropriateness of the three-factor structure of the ELQ (i.e. factorial validity). If the analysis revealed adequate and similar fit indices for both ESEM and CFA models, then there would be less advantage in pursuing an ESEM analysis because the ESEM model is less parsimonious than the CFA model.

To examine the adequacy of the model, the following fit indices were considered: the χ^2/ degrees of freedom (df) statistic, the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) and the WRMR value. According to rules of thumb, good model fit can be stated when CFI and TLI are greater than roughly 0.90, and the RMSEA is smaller than 0.05 (Kline, 2011). Moreover, a ratio of <5 for the χ^2/df statistic usually means that the observed data fits well within the theoretical model proposed, whereas a value smaller than 2 means a significant adjustment (Jöreskog and Sörbom, 1993). Finally, WRMR values of $\leqslant 0.90$ are suggestive of good model fit (Bowen and Guo, 2012).

As the ELQ is based on Likert-type items, the data were analyzed as categorical (ordinal) and not as continuous. The Mplus default estimator for categorical data analyses are weighted least squares with mean and variance adjusted.

Step 3: Invariance of the ELQ items across gender. The goal of this analysis was to examine the DIF of the ELQ to determine the presence of gender biases in items. We wanted to verify whether or not the probability of endorsing each item of the ELQ is equal for both men and women showing an equal degree of leadership. Parametric IRT can be used to distinguish qualitative differences in ways of expressing leadership from quantitative differences in men and women. DIF analyses were conducted by means of the iterative hybrid ordinal logistic regression/IRT (available from the Comprehensive R Archive Network at http://Cran.R-project.org/package = lordif; see Choi *et al.*, 2011, for more details). The presence of DIF was based on the likelihood ratio χ^2 statistical significance test (Swaminathan and Rogers, 1990). A *p*-value <0.01 is suggestive of a DIF.

Step 4: SEM. The objective of this step is to examine the convergent validity of the ELQ. To this end, we verified the relations between the ELQ factors and ethical sensitivity as a criterion variable. In order to conduct SEM with latent variables, it is recommended to use multiple indicators for each variable, because scores from multiple indicators tend to be more reliable and valid than those from a single indicator (Kline, 2011). In the current study, each latent variable was defined by the means of two indicators or parcels (Bandalos and Finney, 2001; Nasser and Takahashi, 2003), where each parcel represented a subset of the scale items. To examine the adequacy of the model testing the relationship between the three constructs underlying the ELQ (ethics of care, justice, and critique) and ethical sensitivity, we used the same indices as presented above, with the exception of WRMR, which was replaced by SRMR. According to rules of thumb (Kline, 2011), good model fit can be stated when SRMR is <0.10.

Missing data. From the total of 668 participants, 576 participants completed all items of the questionnaire while 92 participants had at least one missing value on one of the four scales: 1-5 missing values (60.9 percent), 6-10 (14.1 percent), 13-19

(22.8 percent), and 25-26 (2.2 percent). We chose to present our results using multiple imputation (Allison, 2001) since this method takes full advantage of the available data and avoids some of the bias in standard errors and test statistics that can accompany traditional ad hoc methods such as listwise or pairwise deletion or substituting missing values with the variable mean (Peugh and Enders, 2004). Starting from simple random values, imputation proceeds by iterating over the conditionally specified models (van Buuren, 2010). In the current study, we used a fully conditional specification method which has been implemented in R as the multivariate imputation chained equation package (van Buuren and Oudshoorn, 2000). This strategy was used for the DIF, CFA, ESEM, and SEM analyses, but not for the IRT analysis of the three subscales (ethic of care, ethic of critique, and ethic of justice). In effect, using the imputed data to perform an item analysis was based on IRT which uses the "maximum likelihood estimation" method for estimating item parameters and ability levels for examinees without replacing missing values. In fact, as a result of combining information on the examinee's entire pattern of responses as well as the characteristics of each item for estimating item parameters and ability levels for examinees, using imputed data are expendable or non-essential.

The study presented here has three limitations. First, the convergent validity was established via ethical sensitivity alone. It should be useful in future research to measure other relevant constructs, such as idealized influence, interpersonal justice, and informational justice (Mayer *et al.*, 2012). Second, divergent validity should be demonstrated by including other leadership styles that we expected would be non- or negatively related such as passive and autocratic leadership. Third, construct validity could be more stringently tested via a multi-trait/multi-method approach.

Results
Step 1: Item analysis based on IRT
The item discrimination (a_i), the difficulty (b_{ij}), and informativeness (I_{imax}) values for the items of the scale "ethic of care," "ethic of critique," and "ethic of justice" are summarized in Tables I-III, respectively.

Figures 1 and 2 are examples of good and bad discriminating items, respectively. The first curve represents item 26 (ethic of justice subscale) which states: "When I have to resolve an ethical dilemma, I conduct an investigation." Based on our results, this item discriminates very well between educational leaders with low, moderate, and high levels of ethic of justice ($a_i > 1.70$; Baker, 2001) and the curve of this item is steep at all the levels of the latent trait. In contrast, from the subscale "ethic of critique," we deleted item 29: "The decision I make when faced with an ethical dilemma aims at greater good." The low value of the a_i parameter for the Item 29 (a_i between 0.01 and 0.34; Baker, 2001) indicated that this question was not a good indicator of behavior related to "ethic of critique." Considering its ICC, we noted the flat slope at all the levels of the latent trait (see Figure 2). In sum, following a joint analysis of the a_i, b_{ij}, and I_{imax} parameters, we decided to keep all of the items of the subscale "ethic of justice" and to reject one item from the subscale "ethic of critique" (item 29) and three items from the subscale "ethic of care" (items 2, 18, and 22). A fine-grained analysis of the content of these items indicated that they were ambiguous. Thus, at the end of this first step, the ELQ contained 26 items (ethic of care: ten; ethic of critique: eight; ethic of justice: eight).

Step 2: CFA and ESEM
In this second step, we applied the traditional CFA approach and the more recent ESEM approach (Asparouhov and Muthén, 2009; Marsh *et al.*, 2010) to test the

Item	a_i	b_1 (SE)	b_2 (SE)	b_3 (SE)	b_4 (SE)	b_5 (SE)	b_6 (SE)	I_{imax}
1	0.82		−6.61 (1.93)	−5.76 (1.53)	−3.83 (0.97)	−0.82 (0.27)	1.16 (0.36)	0.20
2	0.02		−147.71 (922.89)	−75.69 (472.71)	22.70 (142.53)	78.29 (489.87)	107.49 (672.44)	0.00
3	1.02			−4.37 (1.03)	−2.77 (0.63)	−0.03 (0.15)	1.10 (0.30)	0.29
5	2.80			−2.60 (0.59)	−2.00 (0.44)	−0.76 (0.17)	0.10 (0.08)	2.07
8	1.54		−3.91 (0.91)	−3.69 (0.81)	−2.34 (0.49)	−0.33 (0.12)	0.56 (0.16)	0.69
9	1.21			−4.80 (1.26)	−3.56 (0.84)	−1.17 (0.29)	−0.04 (0.15)	0.40
12	1.97				−1.85 (0.42)	−1.12 (0.26)	−0.39 (0.14)	1.03
13	1.02		−5.41 (1.40)	−4.36 (1.01)	−2.42 (0.54)	−0.48 (0.18)	0.58 (0.21)	0.31
14	2.16			−2.90 (0.64)	−2.24 (0.47)	−0.59 (0.15)	0.41 (0.13)	1.24
15	2.72			−2.61 (0.59)	−2.04 (0.44)	−0.61 (0.15)	0.26 (0.10)	1.95
18	0.46		−11.06 (4.16)	−8.27 (2.95)	−3.51 (1.23)	1.01 (0.49)	3.53 (1.30)	0.06
19	1.59			−2.74 (0.57)	−1.98 (0.40)	−0.11 (0.10)	1.01 (0.23)	0.70
22	0.34	−12.88 (5.74)	−11.22 (4.90)	−6.84 (2.97)	−1.78 (0.85)	3.35 (1.53)	6.13 (2.73)	0.04

Table I.
Parameters estimates of
the items of the ethic
of care dimension

Item	a_i	b_1 (SE)	b_2 (SE)	b_3 (SE)	b_4 (SE)	b_5 (SE)	b_6 (SE)	I_{imax}
6	1.59	-2.78 (0.52)	-2.30 (0.41)	-1.04 (0.20)	0.11 (0.10)	1.11 (0.22)	1.71 (0.33)	0.78
10	0.63			-4.49 (1.27)	-3.08 (0.87)	0.32 (0.26)	2.30 (0.71)	0.12
11	2.11	-3.22 (0.67)	-2.65 (0.49)	-1.58 (0.28)	-0.55 (0.12)	0.74 (0.15)	1.51 (0.28)	1.29
16	2.44		-2.26 (0.42)	-1.74 (0.31)	-0.70 (0.14)	0.43 (0.10)	1.02 (0.20)	1.67
17	1.76	-2.61 (0.48)	-2.18 (0.39)	-1.16 (0.21)	-0.04 (0.09)	1.21 (0.23)	1.93 (0.36)	0.94
23	0.80		-5.13 (1.27)	-3.74 (0.89)	-1.21 (0.33)	1.36 (0.37)	2.79 (0.70)	0.19
28	1.46	-4.05 (0.93)	-4.05 (0.93)	-3.08 (0.63)	-1.38 (0.28)	0.33 (0.12)	1.31 (0.28)	0.61
29	0.27		-12.87 (7.23)	-10.06 (5.61)	-4.10 (2.32)	4.11 (2.40)	9.18 (5.21)	0.02
30	1.16		-6.65 (1.87)	-5.00 (1.28)	-2.49 (0.62)	-0.11 (0.19)	1.41 (0.41)	0.19

Table II.
Parameters estimates of
the items of the ethic
of critique dimension

Item	a_i	b_1 (SE)	b_2 (SE)	b_3 (SE)	b_4 (SE)	b_5 (SE)	b_6 (SE)	I_{imax}
4	0.85		−6.35 (1.78)	−6.35 (1.78)	−4.16 (1.08)	−0.44 (0.20)	1.09 (0.33)	0.20
7	0.64			−5.68 (1.60)	−3.66 (1.01)	0.04 (0.23)	1.71 (0.54)	0.12
20	0.64			−3.92 (1.08)	−2.55 (0.71)	0.69 (0.30)	2.56 (0.75)	0.12
21	1.64	−3.87 (0.86)	−3.19 (0.62)	−1.86 (0.34)	−0.57 (0.14)	0.88 (0.18)	1.68 (0.32)	0.78
24	2.44	−2.54 (0.50)	−2.24 (0.42)	−1.65 (0.30)	−0.99 (0.19)	−0.08 (0.07)	0.47 (0.12)	1.85
25	1.11	−4.38 (0.99)	−3.70 (0.77)	−2.59 (0.53)	−1.26 (0.28)	0.82 (0.22)	1.99 (0.44)	0.39
26	1.78	−3.67 (0.80)	−3.00 (0.58)	−1.69 (0.31)	−0.55 (0.13)	0.64 (0.14)	1.31 (0.25)	0.94
27	0.92	−5.78 (1.49)	−5.40 (1.31)	−3.79 (0.85)	−1.65 (0.38)	0.28 (0.17)	1.30 (0.34)	0.26

Table III.
Parameters estimates of
the items of the ethic
of justice dimension

proposed three-factor structure of the ELQ. It was necessary to conduct the CFA and the ESEM on the same subgroup of participants in order to verify if the analysis revealed adequate and similar fit indices for both ESEM and CFA models.
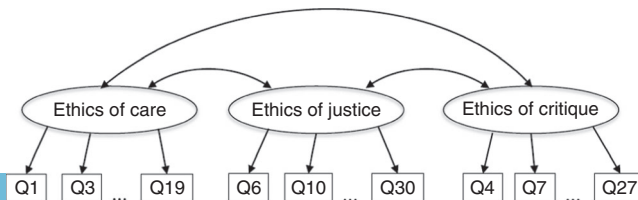
The CFA model (see Figure 3) posits that an item of the ELQ reflects only one latent variable; consequently, the cross-loading is set to zero. Thus, for example, items
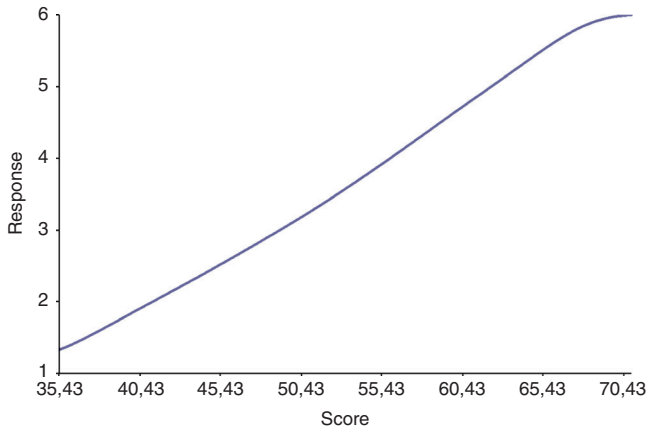
Figure 1.
The item characteristic
curve of item 26
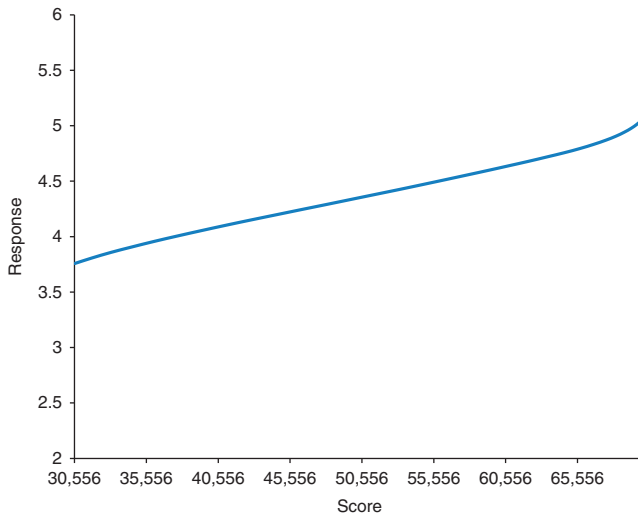
Figure 2.
The item characteristic
curve of item 29

Figure 3.
The three-factor model
of the ELQ

assumed to represent the ethic of justice construct will load solely on this latent variable. As the three ethics are conceptually close, and given that two or more items measuring different latent variables might correlate (Brown, 2006), we also formulated the hypothesis that all the latent variables were correlated. The goodness-of-fit indexes suggests that this model does not fit the data well: $\chi^2/df = 1.69$; CFI = 0.87; TLI = 0.84; RMSEA = 0.06; WRMR = 1.10.

In contrast, when we examined the factorial structure of the ELQ using ESEM (Asparouhov and Muthén, 2009; Marsh *et al.*, 2010), the fit indices suggest an approximate fit of the data: $\chi^2/df = 1.62$; CFI = 0.92; TLI = 0.90; RMSEA = 0.06; WRMR = 0.82. In fact, the results of the ESEM solution (Table IV) indicated that most items had the highest loadings on the factor they are supposed to measure. More specifically, all items designed to measure "ethic of care" had their highest loading on this factor. However, considering the other two factors "ethic of critique" and "ethic of justice," some items had almost equivalent loadings on another factor than the one they were supposed to measure. With regard to "ethic of critique," the loadings of the items varied between 0.270 and 0.798, but one out of the eight items (30)

	CFA			ESEM		
Items of Factor 1: ethic of care						
Q1	0.335	0.000	0.000	0.227	0.088	0.163
Q3	0.500	0.000	0.000	0.575	0.071	0.153
Q5	0.663	0.000	0.000	0.703	0.045	0.142
Q8	0.513	0.000	0.000	0.753	−0.148	0.100
Q9	0.531	0.000	0.000	0.478	0.246	0.127
Q12	0.483	0.000	0.000	0.560	0.115	0.040
Q13	0.306	0.000	0.000	0.245	0.225	−0.020
Q14	0.612	0.000	0.000	0.733	−0.005	0.062
Q15	0.696	0.000	0.000	0.741	0.126	0.024
Q19	0.683	0.000	0.000	0.501	0.237	0.199
Items of Factor 2: ethic of critique						
Q6	0.000	0.336	0.000	0.076	0.471	−0.199
Q10	0.000	0.551	0.000	0.167	0.376	0.236
Q11	0.000	0.624	0.000	−0.080	0.798	0.042
Q16	0.000	0.583	0.000	−0.013	0.798	−0.017
Q17	0.000	0.441	0.000	−0.025	0.613	−0.008
Q23	0.000	0.390	0.000	0.230	0.298	−0.040
Q28	0.000	0.651	0.000	0.048	0.645	0.179
Q30	0.000	0.522	0.000	0.291	0.270	0.184
Items of Factor 3: ethic of justice						
Q4	0.000	0.000	0.361	0.055	0.261	0.180
Q7	0.000	0.000	0.442	−0.046	0.048	0.573
Q20	0.000	0.000	0.469	0.336	0.059	0.229
Q21	0.000	0.000	0.222	−0.135	−0.049	0.492
Q24	0.000	0.000	0.488	0.027	−0.112	0.789
Q25	0.000	0.000	0.678	0.157	0.004	0.646
Q26	0.000	0.000	0.525	−0.093	0.189	0.553
Q27	0.000	0.000	0.536	0.081	0.233	0.385
			Intercorrelations			
Factor 1						
Factor 2	0.559			0.244		
Factor 3	0.590	0.502		0.208	0.303	

Table IV.
Factors loadings for the three intercorrelated-factor CFA and ESEM solutions

had a slightly higher loading on another factor. In the case of "ethic of justice," two out
of the eight items (4, 20) had a slightly higher loading on another factor.

Given these results, and considering that we had a sufficient number of items to
assure the content validity of the ethic of justice dimension, we decided to remove the
items that had much higher loadings on another factor (4, 20, 30). Thus, at the end of
this second step, the ELQ contained 23 items (ethic of care: ten; ethic of critique: seven;
ethic of justice: six).

Step 3: DIF of the ELQ
Parametric IRT with package Lordif (Choi *et al.*, 2011) has been used to test item
invariance across gender. It determines whether men and women with equivalent levels of
leadership endorse items in qualitatively different ways. The results revealed that the
expected item scores do not vary for women and men; in fact, for the 23 items, men's and
women's ICCs overlap to a considerable extent. Non-significant χ^2 values ($p < 0.01$)
indicated that there is no DIF. Considering these results, it appears that the items of the
ELQ are not gender biased (see the final version of the ELQ in the Appendix).

Step 4: SEM
The fit indices for the model testing relationship between the three constructs underlying
the ELQ (ethic of care, justice, and critique) and ethical sensitivity suggested that it does
fit the data well: $\chi^2/df = 2.28$; CFI $= 0.961$; TLI $= 0.934$; RMSEA $= 0.088$ (90 percent
CI $= 0.055$-0.121), and SRMR $= 0.046$. The results showed that ethic of care, justice, and
critique accounted for 16.2 percent of the variance in ethical sensitivity ($p = 0.014$). The
results also indicated that the ethic of critique construct ($\gamma = 0.34$, $p < 0.05$) is significantly
related to ethical sensitivity whereas the other two latent variables (i.e., ethic of care and
ethic of justice) do not predict ethical sensitivity.

Discussion
This study resulted in an instrument with acceptable validity for measuring the presence
of ethical leadership based on three interdependent ethics – justice, critique, and care –
that are linked to a dependent variable, namely, ethical sensitivity. The ELQ stems from
our conception of ethical leadership, rooted in a critical reflective capacity and operating
in three ethical dimensions. The results provide quantitative confirmation of the presence
of the three ethics in the resolution of moral dilemmas. The validity of Starratt's (1991)
model is thus supported qualitatively by our previous studies (Langlois, 1997; Langlois
and Lapointe, 2007, 2010) and quantitatively by the present study.

Furthermore, our results indicated that there is no DIF. Thus, it is reasonable to think
that the ELQ items have the same meaning for both men and women, which means that
gender differences observed when using ELQ will not be due to a faulty instrument but
to differences between women and men. Given the present debate on the issue of
gender-related distinctions between the leadership of women and men, more reliable
results will be available through the use of ELQ in diverse research settings.

With regard to findings which indicate that certain items measuring the ethic of
critique have almost equivalent loadings on the ethic of justice and vice versa, these
results are not surprising given the fact that certain behaviors can be related to more
than one dimension. According to Starratt (1991), the ethic of critique is very close to
the ethic of justice as the use of critical lenses sheds light on injustices in order to attain
greater social justice. Furthermore, the core foundation of the ethic of justice can be
confusing as different schools of thought exist when defining this ethic. In the ELQ,

items pertaining to justice-based moral gestures are partly based on Kohlberg's work. As mentioned earlier, we believe that Kohlberg's Stages 5 and 6 (post-conventional levels) are associated with the ethic of critique, and that transformative and restorative justice both emerge from the ethic of critique.

Another important result concerns the presence of relationships between ethical sensitivity and the ethic of critique. With regard to correlation between ELQ constructs and ethical sensitivity, our results show that the ethic of critique is significantly related to ethical sensitivity while the ethic of care and the ethic of justice do not predict ethical sensitivity. With regard to the ethic of critique, writing about leaders who are inspired by this dimension, Starratt (1991) suggested that "their basic stance is ethical for they are dealing with questions of social justice and human dignity" (p. 189). In a previous, qualitative research project during which a training program was evaluated, we observed the same thing in leaders whose ethical profile indicated a pronounced ethic of critique and who had been able to resolve their ethical dilemmas and initiate meaningful changes within their organization. These leaders were deeply concerned about issues of social justice they witnessed and were ready to take action in order to preserve equity in their schools or districts. As educational organizations and associations become aware of the crucial need for more ethical leaders, they will need to pay particular attention to the ethic of critique as it appears to play a significant role in the development of ethical sensitivity, the ability to discern injustice, and privilege being a sign that one's consciousness and perception of ethical issues is awakened. Such a result deserves attention with regard to principals' pre-service and continuing ethical development as envisioned by Frick *et al.* (2009).

At this point, we want to stress the fact that our objective in developing the ELQ was not to create an instrument for the identification of candidates who would score higher on the three ethics, or on any one of them. Our goal was rather to be able to identify the presence of emergent or confirmed ethical competency among leaders and to support their optimal development through professional training (Langlois and Lapointe, 2010). Although employers might be tempted to use the ELQ in order to select leaders with a particular ethical profile, its intended aim is to support leaders who wish to become more ethically sensitive and, by doing so, attain a more optimal development in all three ethics. Indeed, during the aforementioned experimental training project, we observed a phenomenon where participants experienced a significant readjustment process triggered by the information provided to them through the questionnaire (Langlois and Lapointe, 2010). Upon receiving their pre-training ELQ profile, participants were able to identify the ethic that they had already developed and those that were either not present or were in the process of being acquired. We found this step to be very important as it allowed participants to become aware of their strengths and weaknesses, a recognition which led them to engage in a transformative cycle aimed at improving their ethical awareness, moral judgment, sense of responsibility, and overall professional conduct.

While the current results are promising, our study has some methodological limitations. First, more replications will be needed to fully support the ELQ's validity. For example, while the link between ethical sensitivity and the ethic of critique may suggest the ELQ's convergent validity, to correlate ELQ results with other external criterion or validated available scales would strengthen its validity. Moreover, the focus of the invariance assessed was gender. Given that the instrument may be used in diverse cultural contexts, invariance across cultures would be warranted. Finally, different subgroups of participants were used at each step of the psychometric data analysis.

However, these subgroups were drawn from one sample. Our findings should be replicated in studies using larger, more representative samples that include, for instance, school principals from ethnic minorities as well those coming from diverse countries.

Conclusion

Scholars in different fields, including education, business, and public administration, have stressed the urgent need for a better understanding of ethical leadership and its main components, in particular ethical sensitivity and conduct. Many have recognized the absence of validated research instruments that would help grasp how ethical leadership is enacted and how to define the concept more accurately. Results presented in this paper are meant to help fill the void at a time where there is a vital need for more ethical attitudes and for greater integrity in education and in the workplace in general.

Note

1. In order to emphasize the pioneering nature of Shapiro and Stefkovich's work as well as of Strike *et al.*, we have identified the years their respective books were first published.

References

Allison, P.D. (2001), *Missing Data*, Sage, Thousand Oaks, CA.

Asparouhov, T. and Muthén, B. (2009), "Exploratory structural equation modeling", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 16 No. 3, pp. 397-438.

Baker, F.B. (2001), *The Basics of Item Response Theory*, ERIC Clearinghouse on Assessment and Evaluation, College Park, MD.

Bandalos, D.L. and Finney, S.J. (2001), "Item parceling issues in structural equation modeling", in Marcoulides, G.A. and Schumacker, R.E. (Eds), *Advanced Structural Equation Modeling: New Developments and Techniques*, Lawrence Erlbaum Associates Inc, Mahwah, NJ, pp. 269-296.

Bates, R.J. (1982), "Towards a critical practice of educational administration", paper presented at the Annual meeting of the American Educational Research Association, New York, NY, April 1-5.

Beck, L.G. (1994), *Reclaiming Educational Administration as a Caring Profession*, Teachers College Press, New York, NY.

Begley, P.T. and Johansson, O. (2003), *The Ethical Dimensions of School Leadership*, Falmer Press, New York, NY.

Bowen, K.N. and Guo, S. (2012), *Structural Equation Modeling*, Oxford University Press, New York, NY.

Brown, L.M., Argyris, D., Attanucci, J., Bardige, B., Gilligan, C., Johnston, D.K. and Wilcox, D. (1988), *A Guide to Reading Narratives of Conflict and Choice for Self and Relational Voice (Monograph No. 1)*, Project on the Psychology of Women and the Development of Girls, Harvard Graduate School of Education, Cambridge, MA.

Brown, M., Treviño, J.K. and Harrisson, D. (2005), "Ethical leadership: social learning perspective for construct development and testing", *Organizational Behavior and Human Decision Processes*, Vol. 97 No. 2, pp. 117-130.

Brown, T.A. (2006), *Confirmatory Factor Analysis for Applied Research*, The Guilford Press, New York, NY.

Brunner, C. (1998), "Can power support an ethic of care? An examination of professional practices of women superintendents", *Journal for a Just and Caring Education*, Vol. 4 No. 2, pp. 160-182.

Butterfield, K.D., Trevino, L.K. and Weaver, G. (2000), "Moral awareness in business organizations: influences of issue-related and social context factors", *Human Relations*, Vol. 53 No. 7, pp. 981-1018.

Choi, S.W., Gibbons, L.E. and Crane, P.K. (2011), "Lordif: an R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations", *Journal of Statistical Software*, Vol. 39 No. 8, pp. 1-30.

Clarkeburn, H. (2002), "A test for ethical sensitivity in science", *Journal of Moral Education*, Vol. 31 No. 4, pp. 439-453.

Colvin, G. (2003), "Corporate crooks are not all created equal", *Fortune*, Vol. 148 No. 9, p. 64.

Cooper, T.L. (2006), *The Responsible Administrator: An Approach to Ethics for the Administrative Role*, 5th ed., Jossey-Bass, San Francisco, CA.

Cranston, N., Ehrich, L.C. and Kimber, M. (2005), "Ethical dilemmas: the 'bread and butter' of educational leaders lives", *Journal of Educational Administration*, Vol. 44 No. 2, pp. 106-121.

Crowson, R. (1989), "Managerial ethic in educational administration: the rational choice approach", *Urban Education*, Vol. 23 No. 4, pp. 412-435.

Damasio, A. (1999), *Le sentiment même de soi : corps, émotions, conscience (The Very Feeling of One's Self: Body, Emotions, Conscience)*, Odile Jacob, Paris.

Edmonson, S., Fisher, A. and Polnick, B. (2003), "Portrait of an ethical administrator", paper presented at the Annual Conference of the American Association of School Administrators, New Orleans, LA, April 21-25.

Embretson, S.E. and Reise, S.P. (2000), *Item Response Theory for Psychologists*, Lawrence Erlbaum, Mahwah, NJ.

Enomoto, E. and Kramer, B.H. (2007), *Leading Through the Quagmire: Ethical Foundations, Critical Methods, and Practical Applications for School Leadership*, Rowman & Littlefield Education, Lanham, MD.

Eyal, O., Berkovitch, I. and Schwartz, T. (2011), "Making the right choices: ethical judgments among educational leaders", *Journal of Educational Administration*, Vol. 49 No. 4, pp. 396-413.

Foster, W. (1986), *Paradigms and Promises: New Approaches to Educational Administration*, Prometheus, Buffalo, NY.

Frick, W.C., Polizzi, J.A. and Frick, J.E. (2009), "Aspiring to a continuous learning ethic: building authentic learning communities for faculty and administration", *Education Leadership and Administration*, Vol. 21, pp. 7-26.

Gilligan, C. (1982), *In a Different Voice: Psychological Theory and Women's Development*, Harvard University Press, Cambridge, MA.

Greenfield, W. (1987), "Moral imagination and value leadership in school", paper presented at the Annual Convention of the American Educational Research Association, Washington, DC, April 9-13.

Halpin, A.W. (1970), *Administrative Theory in Education*, Macmillan, New York, NY.

Hatcher, W.S. (1998), *Love, Power and Justice: The Dynamics of Authentic Morality*, Baha'i Publishing Trust, Wilmette, IL.

Hodgkinson, C. (1978), *The Philosophy of Leadership*, Blackwell, Oxford.

Jöreskog, K. and Sörbom, D. (1993), *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*, Scientific Software International Inc, Chicago, IL.

Kalshoven, K. (2010), "Ethical leadership: through the eyes of employees", doctoral dissertation, University of Amsterdam, Amsterdam, available at: http://dare.uva.nl/record/337849

Kalshoven, K., Den Hartog, D.N. and De Hoogh, A.H.B. (2011), "Ethical leadership at work (ELW): development and validation of a multidimensional measure", *The Leadership Quarterly*, Vol. 22 No. 1, pp. 51-69 (accessed May 12, 2011).

Kirby, P.C., Paradise, L.V. and Protti, R. (1992), "Ethical reasoning of educational administrators: structuring inquiry around the problems of practice", *Journal of Educational Administration*, Vol. 30 No. 4, pp. 25-32.

Kline, R.B. (2011), *Principles and Practices of Structural Equation Modelling*, 3rd ed., Guilford Press, New York, NY.

Kohlberg, L. (1972), "Development as the aim of education", *Harvard Educational Review*, Vol. 42 No. 4, pp. 448-495.

Kohlberg, L. (1981), *Essays on Moral Development I: The Philosophy of Moral Development*, Harper & Row, San Francisco, CA.

Langlois, L. (1997), "Relever les défis de la gestion scolaire d'après un modèle de leadership éthique : une étude de cas (Taking up challenges as educational administrators based on an ethical leadership model)", doctoral dissertation, Université Laval, Québec.

Langlois, L. (2004), "Responding ethically: complex decision-making by school district superintendents", *International Studies Educational Administration Management*, Vol. 32 No. 2, pp. 83-99.

Langlois, L. and Lapointe, C. (2007), "Ethical leadership in Canadian school organizations: tensions and possibilities", *Educational Management, Administration, and Leadership*, Vol. 35 No. 2, pp. 247-260.

Langlois, L. and Lapointe, C. (2009), "Navigating by sight in a world without lighthouses: the challenge of being ethically professional", paper presented at the 14th Annual Values and Leadership Conference, Pennsylvania State University, State College, PA, April 15-19.

Langlois, L. and Lapointe, C. (2010), "Can ethic be learned? Results from a three-year action-research project", *Journal of Educational Administration*, Vol. 48 No. 2, pp. 147-163.

Lavorata, L. (2005), "Lien entre climat éthique et comportement éthique du vendeur : rôle déterminant du management commercial (Links between ethical climate and sellers' ethical conduct: the determining role of business management)", *Actes du 21ᵉcongrès de l'AFM, Nancy*.

Liedtka, J. (1992), "Exploring ethical issues using personal interviews", *Business Ethics Quarterly*, Vol. 2 No. 2, pp. 161-181.

Marsh, H.W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A.J.S., Trautwein, U. and Nagengast, B. (2010), "A new look at the Big Five factor structure through exploratory structural equation modeling", *Psychological Assessment*, Vol. 22 No. 3, pp. 471-491.

Marsh, H.W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A.J.S. and Trautwein, U. (2009), "Exploratory structural equation modelling, integrating CFA and EFA: application to students' evaluations of university teaching", *Structural Equation Modeling*, Vol. 16 No. 3, pp. 439-476.

Marshall, C., Paterson, J.A., Rogers, D.L. and Steele, J.R (1993), "Caring as career: an alternative model for educational administration", paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA, April 12-16.

Maxcy, S.J. (2002), *Ethical School Leadership*, Scarecrow Press, Lanham, MD.

Mayer, D.M., Aquino, K., Greenbaum, R.L. and Kuenzi, M. (2012), "Who displays ethical leadership, and why does it matter? An examination of antecedents and consequences of ethical leadership", *Academy of Management Journal*, Vol. 55 No. 1, pp. 151-171.

Mehta, S. (2003), "MCI: Is being good, good enough? Mike Capellas has his plate full cleaning up MCI. That's just a start", *Fortune*, Vol. 148 No. 9, p. 117.

Miles, M.B. and Huberman, M.A. (2003), *Analyse des données qualitatives (Qualitative Data Analysis)*, 2nd ed., De Boeck, Paris.

Muthén, L.K. and Muthén, B.O. (1998/2010), *Mplus User's Guide*, 6th ed., Muthén & Muthén, Los Angeles, CA.

Nasser, F. and Takahashi, T. (2003), "The effect of using item parcels on ad hoc goodness-of-fit indexes in confirmatory factor analysis: an example using Sarason's reactions to tests", *Applied Measurement in Education*, Vol. 16 No. 1, pp. 75-97.

Noddings, N. (1984), *Caring: A Feminine Approach to Ethic and Moral Education*, University of California Press, Berkeley, CA.

Paillé, P. and Muchielli, A. (2003), *L'analyse qualitative en sciences humaines et sociales (Qualitative Analysis in Human and Social Sciences)*, Armand Colin, Paris.

Peugh, J.L. and Enders, C.K. (2004), "Missing data in educational research: a review of reporting practices and suggestions for improvement", *Review of Educational Research*, Vol. 74 No. 4, pp. 525-556.

Poitras, S.-C., Guay, F. and Ratelle, C.F. (2012), "Using the self-directed search in research: selecting a representative pool of items to measure vocational interests", *Journal of Career Development*, Vol. 39 No. 2, pp. 186-207.

Reeve, B.B. and Fayers, P. (2005), "Applying item response theory modelling for evaluating questionnaire item and scale properties", in Fayers, P. and Hays, R.D. (Eds), *Assessing Quality of Life in Clinical Trials: Methods and Practice*, 2nd ed., Oxford University Press, New York, NY, pp. 55-73.

Rest, J.R. (1986), *Moral Development: Advances in Research and Theory*, Praeger, New York, NY.

Riggio, R.E., Zhu, W., Reina, C. and Maroosis, J.A. (2010), "Virtue-based measurement of ethical leadership: the leadership virtues questionnaire", *Consulting Psychology Journal: Practice and Research*, Vol. 62 No. 4, pp. 235-250.

Samejima, F. (1969), "Estimation of ability using a response pattern of graded scores", *Psychometrika Monograph Supplement*, Vol. 34 No. 4, pp. 1-97.

Samejima, F. (1997), "Graded response model", in van der Linden, W.J. and Hambleton, R.K. (Eds), *Handbook of Modern Item Response Theory*, Springer-Verlag, New York, NY, pp. 85-100.

Starratt, R.J. (1991), "Building an ethical school: a theory for practice in educational leadership", *Educational Administration Quarterly*, Vol. 27 No. 2, pp. 185-202.

Starratt, R.J. (2004), *Ethical Leadership*, Jossey-Bass, San Francisco, CA.

Starratt, R.J. (2012), *Cultivating an Ethical School*, Routledge, New York, NY.

Stefkovich, J.A. and Shapiro, J.P. (1995), "Personal and professional ethics for educational administrators: non-traditional and pedagogical implications", *Review Journal of Philosophy and Social Science*, Vol. 20 Nos 1-2, pp. 157-186.

Stiggins, R. (2004), "New assessment beliefs for a new school mission", *Phi Delta Kappan*, Vol. 86 No. 1, pp. 22-27.

Strike, K.A., Haller, E.J. and Soltis, J.F. (1998), *The Ethic of School Administration*, Teacher College Press, New York, NY.

Swaminathan, H. and Rogers, H.J. (1990), "Detecting differential item functioning using logistic regression procedures", *Journal of Educational Measurement*, Vol. 27 No. 4, pp. 361-370.

Treviño, L.K. and Weaver, G.R. (1998), "Punishment in organizations: descriptive and normative perspectives", in Schminke, M. (Ed), *Managerial Ethic: Moral management of People and Processes*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 99-114.

Tuana, N. (2007), "Conceptualizing moral literacy", *Journal of Educational Administration*, Vol. 45 No. 4, pp. 364-378.

Valois, P., Houssemand, C., Germain, S. and Abdous, B. (2011), "An open source tool to verify the psychometric properties of an evaluation instrument", *Procedia-Social and Behavioral Sciences*, Vol. 15, pp. 552-556.

van Buuren, S. (2010), "Item imputation without specifying scale structure", *Methodology: A European Journal of Research Methods for the Behavioral and Social Sciences*, Vol. 6 No. 1, pp. 31-36.

van Buuren, S. and Oudshoorn, C.G.M. (2000), *Multivariate Imputation by Chained Equations: MICE V1.0 User's Manual*, TNO Report No. PG/VGZ/00.038 TNO Prevention and Health, Leiden.

Walumbwa, F.O., Avolio, B.J., Gardner, W.L., Wernsing, T.S. and Peterson, S.J. (2008), "Authentic leadership: development and validation of a theory-based measure", *Journal of Management*, Vol. 34 No. 1, pp. 89-126.

Yukl, G., Mahsud, R., Hassan, S. and Prussia, G.E. (2013), "An improved measure of ethical leadership", *Journal of Leadership & Organizational Studies*, Vol. 20 No. 1, pp. 38-48.

Zheng, X., Zhu, W., Yu, H., Zhang, X. and Zhang, L. (2011), "Ethical leadership in Chinese organizations: developing a scale", *Frontiers of Business Research in China*, Vol. 5 No. 1, pp. 179-198.

Appendix. Ethical Leadership Questionnaire

Definition: an ethical dilemma is a situation that will often involve an apparent conflict between values in which to obey one would result in transgressing another. The crucial features of a moral dilemma are these: the agent is required to do each of two (or more) actions; the agent can do each of the actions; but the agent cannot do both (or all) of the actions.

Referring to the scale below, circle the number of your choice

Never	Rarely	Sometimes	Often	Very often	Always	N/A (not applicable)
1	2	3	4	5	6	X

When I reflect on the way I act at work, I can see that…

1.	I establish trust in my relationships with others	1	2	3	4	5	6	X
2.	I try to ensure harmony in the organization	1	2	3	4	5	6	X
3.	I don't tolerate arrogance	1	2	3	4	5	6	X
4.	I follow procedures and rules	1	2	3	4	5	6	X
5.	I try to preserve everyone's safety and well-being	1	2	3	4	5	6	X
6.	I try to make people aware that some situations disproportionately privilege some groups	1	2	3	4	5	6	X
7.	I speak out against unfair practices	1	2	3	4	5	6	X
8.	I seek to protect each individual's dignity	1	2	3	4	5	6	X
9.	I expect people to make mistakes (it's human nature)	1	2	3	4	5	6	X
10.	I speak out against injustice	1	2	3	4	5	6	X
11.	I am concerned when individuals or groups have advantages compared to others	1	2	3	4	5	6	X

When I have to resolve an ethical dilemma…

12.	I check the legal and regulatory clauses that might apply	1	2	3	4	5	6	X
13.	I check my organiation's unwritten rules	1	2	3	4	5	6	X
14.	I conduct an investigation	1	2	3	4	5	6	X
15.	I sanction mistakes in proportion to their seriousness	1	2	3	4	5	6	X
16.	I try to oppose injustice	1	2	3	4	5	6	X
17.	I take time to listen to the people involved in a situation	1	2	3	4	5	6	X
18.	I seek to preserve bonds and harmony within the organization	1	2	3	4	5	6	X
19.	I avoid hurting people's feelings by maintaining their dignity	1	2	3	4	5	6	X
20.	I pay attention to individuals	1	2	3	4	5	6	X
21.	I promote dialogue about contentious issues	1	2	3	4	5	6	X

My decision in the resolution of an ethical dilemma is based on…

22.	the statutory and legal framework.	1	2	3	4	5	6	X
23.	greater social justice.	1	2	3	4	5	6	X

Care = mean of items 1, 2, 5, 8, 9, 17, 18, 19, 20, 21
Critique = mean of items 3, 6, 7, 10, 11, 16, 23
Justice = mean of items 4, 12, 13, 14, 15, 22

Corresponding author
Dr Lyse Langlois can be contacted at: lyse.langlois@rlt.ulaval.ca