Learning and Individual Differences 41 (2015) 1–13

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

Motivation for PhD studies: Scale development and validation☆

David Litalien [a,*], Frédéric Guay [b], Alexandre J.S. Morin [a]

[a] Australian Catholic University, Australia
[b] Université Laval, Canada

ARTICLE INFO

Article history:
Received 2 July 2014
Received in revised form 8 May 2015
Accepted 23 May 2015

Keywords:
PhD studies
Academic motivation scale
Self-determination theory

ABSTRACT

In Canada and the United States, doctoral attrition rates are estimated to vary from 40% to 60%. Motivation has been proposed as a determinant of doctoral degree completion. The purpose of this study was to develop and validate a scale based on self-determination theory, to assess five types of regulation (intrinsic, integrated, identified, introjected, and external) toward PhD studies. Based on two samples (N = 244, N = 1060), this study involved five steps: (1) item development, (2) factor validation, (3) reliability assessment, (4) convergent and discriminant validity assessment, and (5) measurement invariance testing. Findings from both samples were similar, supporting a five-factor first-order structure and a two-factor higher-order structure, scale reliability, and convergent and discriminant validity as shown by correlations among motivation subscales and correlations between each subscale and various outcomes. Additionally, complete measurement invariance was supported across gender, citizenship status, program type, age, and program progression.

1. Introduction

In OECD (Organization for Economic Cooperation and Development) countries, the number of doctoral degrees awarded grew by 40% in only eight years (from 140,000 in 1998 to 200,000 in 2006; Auriol, 2010). Even in Canada and the United States, where a lower increase had been expected, enrollment in doctoral programs rose by 57% and 64%, respectively, between 1998 and 2010 (OECD, 2013). This growing interest in doctoral studies is partly explained by perceived individual benefits, such as higher income, wider employment opportunities, better working conditions, and increased professional and personal mobility (Auriol, 2010; HRSDC, 2006, Statistics Canada and HRSDC, 2009). Moreover, through their research, they produce and disseminate knowledge, develop innovations, and facilitate social and economic development (AUCC, 2009; Bloom, Hartley, & Rosovsky, 2006; Wendler et al., 2012). However, despite the intensified enrollment and associated benefits, doctoral attrition rates, which fluctuate widely across programs, remain high in North America, estimated at 40% to 50% (MERS, 2013; Nettles & Millett, 2006). Even among rigorously selected students receiving prestigious fellowships, dropout rates can be as high as 25% (Wendler et al., 2010).

Although some students may quit school for practical reasons (e.g., attractive job opportunities, family issues), the consequences for others, as well as universities and society, are unfortunate. Students who dropout find fewer employment opportunities, and their self-esteem can be negatively affected (Lovitts, 2001; Statistics Canada and HRSDC, 2003). Moreover, the substantial time and energy they invested in their studies could have been directed to other areas of their personal and professional lives. For universities, doctoral attrition reduces resources while incurring costs for faculty members having invested considerable time in research projects that will remain incomplete. For society, non-completion of doctoral studies results in lower productivity and competitiveness compared to other countries (Wendler et al., 2010, 2012).

Regardless of the education level, motivation has become a central concept in the understanding of academic persistence and achievement (Pintrich, 2003) and could be particularly important in helping PhD students achieve their goals. Compared to other education levels, doctoral studies are conducted in less structured environments, demand greater independence, involve heavier workloads (e.g., conducting research, publishing results), and encompass more complex tasks. Furthermore, PhD students must invest a considerable amount of time in their studies.

In previous studies, motivation has been proposed as a key construct to explain why some students successfully complete their PhD studies while others do not (see Bair & Haworth, 2005 and Reamer, 1990, for a review; Ivankova & Stick, 2007; Lovitts, 2001). Both in surveys and interviews, students commonly report motivation (or lack thereof) as a reason for leaving or persisting in their program. Most studies

☆ Authors' note: This research was supported a grant from the Social Sciences and Humanities Research Council of canada (SSHRC) and by the Canada Research Chair on Motivation and Academic Success. The third author's work in the preparation of this article was supported by a research grant from the Australian Research Council (DP130102713; DP140101559) and the first author's work was supported by a research grant from the Quebec Fund for Research, Society and Culture. A substantial part of this paper was prepared while the first author was completing his Ph.D. studies at Université Laval (Québec, Canada).

* Corresponding author at: 25A Barker Road, Locked Bag 2002, Strathfield, NSW 2135, Australia.
E-mail address: David.Litalien@acu.edu.au (D. Litalien).

http://dx.doi.org/10.1016/j.lindif.2015.05.006
1041-6080/

interested in doctoral students' motivation are qualitative (e.g., Austin, 2002; Cardona, 2013; Jablonski, 2001; Kärner, Kukemelk, & Herdlein, 2005; O'Meara, Knudsen, & Jones, 2013; see also Bair & Haworth, 2005; Reamer, 1990, for a review) and underscore the relevance of this construct to persistence and success. Diverse motives have emerged from these studies (e.g., intrinsic versus extrinsic reasons, Cardona, 2013; Ivankova & Stick, 2007; personal and professional reasons, Hoskins & Goldberg, 2005; unwillingness to experience failure, Clewell, 1987; commitment, O'Meara et al., 2013), suggesting that motivation to pursue PhD studies is multifaceted. Although insightful, qualitative studies are based on small samples and on a specific discipline or population (e.g., African–American; King & Chepyator-Thomson, 1996), thereby limiting the generalizability of the findings.

Some quantitative studies have also looked at doctoral students' motivation. Unfortunately, most of these studies have neglected to consider the multidimensionality of this construct. It has thus been conceptualized as a single dimension, assessed with a self-report scale (e.g., self-motivation, Ivankova & Stick, 2007), a single item embedded in a list of potential reasons for noncompleters' departures (Lovitts, 2001), or a single item asking students to evaluate their level of motivation compared to their peers (Pauley, Cunningham, & Toth, 1999). As an exception, Anderson and Swazey (1998) asked 2,000 students to assess the importance of various reasons for undertaking doctoral studies. Several reasons were endorsed, such as the desire to gain knowledge in a specific field, conduct research, teach in higher education, and get a well-paying job.

Despite these efforts, the assessment of motivation for PhD studies has rarely been based on a valid theoretical framework. According to Cardona (2013), the development of appropriate conceptual frameworks for understanding motivation at this academic level has been hindered by the complexity of this multifaceted construct and the assumption that students' experiences at this level are specific to their academic discipline. We believe that, in order to gain deeper insight into students' reasons for pursuing a doctoral degree, further research needs to adopt a multidimensional perspective based on a well-established theoretical framework.

One motivation theory that has demonstrated its value and validity in the context of education is self-determination theory (SDT; Ryan & Deci, 2009). A key proposition of SDT is that more internalized regulations (i.e., the person fully endorses the behavior) produce more positive outcomes than less internalized forms of regulations (i.e., the behavior is performed due to internal pressures or external reasons). This perspective has been well supported in primary, secondary and college students (see Guay, Ratelle, & Chanal, 2008), and appears to be well-suited for understanding motivation and persistence in doctoral students as well. Although the distinction between more or less internalized types of regulation has rarely been applied to graduate students, previous studies have found interesting results (see Ahmed & Bruinsma, 2006; Losier, 1994), to which we devote more attention below. However, these studies also included weaknesses. The goal of this study was therefore to develop and validate an SDT-based scale to assess motivation for PhD studies, called the Motivation for PhD Studies scale (MPhD).

1.1. Self-determination theory

SDT proposes that various types of motivation regulate human behavior (Deci & Ryan, 1985, 2012). Intrinsic regulation refers to performing an activity for its own sake, for interest and enjoyment. In contrast, extrinsic motivation refers to engaging in an activity as a means to an end that is separate from the activity itself (Deci & Ryan, 2012). To better account for the motivational process, Deci and Ryan (1985, 2012) suggested that extrinsic motivation consists of four types of regulation that reflect various levels of self-determination. From low to high self-determination, these are external regulation, introjected regulation, identified regulation, and integrated regulation (Deci & Ryan, 1985, 2012). External regulation occurs when an individual adopts a behavior to obtain a reward or to avoid punishment. Introjected regulation occurs when the individual is driven by internal pressure either to pursue self-aggrandizement and contingent self-worth or to avoid guilt and shame. When behaviors are more internalized, accepted and valued, as in identified regulation, individuals consider their behaviors to be important in themselves. Integrated regulation is the most autonomous form of regulation, occurring when behaviors are congruent with the personal goals, values and needs that constitute the self. Whereas identified and integrated regulations underlie a greater sense of autonomy, they remain extrinsic, as the desired outcome remains separable from the activity itself.

According to SDT, these five types of regulation can be situated along an autonomy continuum encompassing, in order, intrinsic, integrated, identified, introjected and external forms of regulation, where intrinsic regulation is the most autonomous type of motivation and external regulation, the least autonomous (Deci & Ryan, 1985, 2012). Because these regulations are aligned on a continuum, they are expected to show a simplex correlation pattern, with stronger positive correlations between adjacent forms of regulation than among more distal forms. For example, intrinsic and integrated regulations should be positively correlated, whereas intrinsic regulation should be more weakly (and potentially negatively) correlated with external regulation.

In line with this continuum, SDT distinguishes two broader categories of motivation: autonomous (including intrinsic, integrated and identified regulations) and controlled (including external and introjected regulations). Autonomous motivation appears to be associated with positive outcomes, such as intention to persist (Black & Deci, 2000; Vallerand, Fortier, & Guay, 1997), performance in course-related activities (Boiché, Sarrazin, Grouzet, Pelletier, & Chanal, 2008), and subjective well-being (Litalien, Lüdtke, Parker, & Trautwein, 2013), whereas controlled motivation appears to be associated with negative outcomes, such as rote learning (Vansteenkiste, Simons, Lens, Sheldon, & Deci, 2004), anxiety (Ryan & Connell, 1989) and lower positive affect (Gillet, Lafrenière, Vallerand, Huart, & Fouquereau, 2014).

1.2. Assessing motivation for PhD studies from an SDT perspective

In the past 30 years, various scales have been developed to assess the different types of regulation proposed by SDT across a wide variety of contexts (e.g., sports, education, and work; Deci & Ryan, 2000). However, few SDT-based studies have assessed motivation in doctoral students. To our knowledge, only Ahmed and Bruinsma (2006) and Losier (1994) have investigated motivation in doctoral and master's students from the SDT perspective. However, they used a slightly adapted version of the Academic Motivation Scale (AMS; Vallerand, Blais, Brière, & Pelletier, 1989; Vallerand et al., 1992), which was developed mainly for high school and college students. Only minor changes were made to the wording (e.g., “high school” was replaced by “graduate studies”), and no questions were added to address students' doctoral dissertations, research skills development, or advisors, all of which are considered key components of doctoral programs. Moreover, these modified versions were never empirically validated. Changes made to the AMS and other instruments to measure doctoral motivation could result in limited explanatory or predictive value, because the items may have little relevance to doctoral studies and could lead to ambiguity about what is being measured (e.g., highly different interpretations of items).

Furthermore, the AMS was originally designed for younger students, and does not include items to assess integrated regulation, which occurs in a more advanced stage of psychological development when the person's identity has been formed (Ratelle, Guay, Vallerand, Larose, & Senécal, 2007). This type of regulation could be particularly relevant for graduate students, who usually have to juggle a number of roles (e.g., worker, spouse, and parent) that might interfere with their studies. Students who fully integrate their behaviors might be more likely to overcome obstacles to academic success (see McLachlan, Spray, & Hagger, 2011, for an example in the field of physical activity).

Although not directly related to PhD studies, Deemer, Martens, and Buboltz (2010) developed a scale inspired by SDT and achievement motivation theory (Elliot, 1997) to measure research motivation in graduate students. However, this scale assesses only the classic intrinsic and extrinsic motivation dichotomy, without considering the more complex types of extrinsic regulations (i.e., integrated, identified, and introjected). For all these reasons, a new scale to assess types of regulation toward PhD studies appears to be needed.

1.3. The present study

The goal of this study was to develop and validate an SDT-based scale to assess motivation for PhD studies. This study relied on two independent samples and involved five steps: (1) item development, (2) factor validation, (3) reliability assessment, (4) convergent and discriminant validity assessment, and (5) measurement invariance testing. We expected our scale to successfully evaluate the five types of motivation and the two broader categories of motivation (autonomous and controlled) to emerge as higher-order factors within a second-order structure.

Convergent and discriminant validity were first assessed by examining the correlations among subscales, with the expectation that the correlations among regulation types would corroborate the simplex correlation pattern theorized by SDT. Second, they were assessed by examining the correlations between the subscales and a variety of positive and negative outcomes. Here, our expectation was that more autonomous forms of regulation (intrinsic, integrated, and identified) would be positively related to positive outcomes (e.g., positive affect, satisfaction, and intention to pursue postdoctoral research) and negatively associated with less desirable outcomes (e.g., dropout intention), while more controlled forms of regulation (introjected and external) would show the opposite pattern. Additionally, intrinsic regulation should have resulted in more positive outcomes compared to integrated regulation, which in turn should have resulted in more positive consequences compared to identified regulation. External regulation should also have been more strongly associated with negative consequences compared to introjected regulation.

The outcome variables were selected for their relevance to SDT, PhD studies, or both. According to SDT, a central consequence of being autonomously motivated is well-being (Ryan & Deci, 2000). Diener, Suh, Lucas, and Smith (1999) conceptualized subjective well-being as a multidimensional construct that contains pleasant affect, the relative absence of unpleasant affect, overall life satisfaction, and satisfaction with certain life domains (e.g., studies). Litalien et al. (2013) showed that autonomous motivation positively predicted these various indicators of well-being among young adults. Test anxiety is also a common problem among university students that has been negatively associated with performance and success (Hembree, 1988). Doctoral students face various assessments (e.g., courses and candidacy examination) that could generate stress and anxiety. Among adult learners, test anxiety has been positively associated with controlled motivation and negatively associated with autonomous motivation (Vansteenkiste, Zhou, Lens, & Soenens, 2005). Additionally, lower levels of autonomous motivation have been associated with dropout intention in high school (Vallerand et al., 1997). Postdoctoral intention, thesis difficulties and perceived academic performance were also assessed as relevant indicators of PhD study experiences.

To verify the generalizability of the results of the MPhD scale and the suitability of this scale for students of diverse backgrounds, we also tested the measurement invariance of the five-factor structure and the higher-order structure across both samples and across various relevant subgroups of PhD students. These subgroups were formed on the basis of gender, citizenship status, program type, age, and study progression.

2. Method

2.1. Participants and procedures

2.1.1. Sample 1

In March 2011, an email invitation was sent to every doctorate student at a large French-language university in Canada using a general email list (listserve; $N = 2319$, including a small proportion of non-PhD students). Although all students on this list were contacted, only PhD students were asked to complete a 154-item online questionnaire (using CallWeb), which took about 35 min to fill out. A total of 339 PhD students participated voluntarily, with no financial incentive. Participation in the study was confidential. Ninety-five participants did not complete any of the MPhD items. We ran analyses on the sample of 244 individuals who completed at least part of the scale (approximately 11% of the eligible sample). This subsample was nearly free of missing data; only three participants omitted to answer three items or fewer. Mean age was 30.7 years ($SD = 6.2$) and 63.9% were female. Participants differed in their program progression and had completed an average of 6.7 trimesters ($SD = 4.7$; a normal study year includes three trimesters). More than half (54.7%) were in natural sciences programs (45.3% in human sciences). With respect to citizenship, 70% were Canadian, 7% were permanent residents, and 23% held temporary visas.

2.1.2. Sample 2

Several steps were taken to ensure that Sample 2 was sufficiently large. First, in October 2011, an email was sent to all PhD students enrolled in a French-speaking Canadian university ($N = 2266$) to invite them to participate in a study on determinants of doctoral persistence. We asked them to complete an online questionnaire lasting about 40 min. We subsequently used various strategies to remind students that their participation was important: an email to faculty members asking for their help in recruiting, two personalized emails, phone calls and, finally, a letter. Students who agreed to complete the questionnaire were eligible for a draw of two iPads. A total of 1,060 PhD students participated in the study (48% of the eligible sample). Their mean age was 31.9 years ($SD = 8.1$), and 52.1% were female. The participants, who had completed an average of 7.1 trimesters ($SD = 5.5$), were enrolled in 71 programs and 17 faculties. Half the participants were in natural sciences programs (50.7%) and the other half, in human sciences (49.3%). With respect to citizenship, 67.4% were Canadian, 9.1% were permanent residents, and 23.5% held temporary visas.

2.2. Developing the motivation for PhD studies scale (MPhD)

The MPhD scale was developed to assess each type of regulation toward PhD studies proposed within the SDT framework. A group of experts on SDT (i.e., two professors, each with over 10 publications based on SDT, and two PhD students who did their empirical master thesis using SDT and who were currently focusing on this theoretical framework in their PhD thesis) first developed a pool of items to assess possible reasons to persist in doctoral studies according to the five SDT regulation types. Seventeen PhD students were then invited to test the scale. To ensure that the items captured the main reasons for academic persistence, students were first asked to write down 10 reasons for persisting in their doctoral studies. To assess the face validity of the items generated by the expert committee, they were then asked to rate the items on their relevance to motivation for PhD studies and on the clarity of the wording (using five-point scales). Six items that the students considered unclear or irrelevant were reworded or replaced. In addition, reasons that were frequently cited by the students but not comprised in the initial item pool were included, resulting in a 25-item scale. This new pool of items was then reexamined by the expert group. Their task was to ensure that every retained item had a sufficient level of face and content validity. Various studies have used these strategies, especially in the health domain (e.g., Broder, McGrath, & Cisneros, 2007; Redsell, Lennon, Hastings, & Fraser, 2004). Retained items had to comply with three face-validity criteria: (1) relevance for assessing PhD students' motivation; (2) applicability to most PhD students, regardless of their program or progression; and (3) wording clarity and concision. Regarding content validity, the items had to meet two criteria: (1) relevance to the assessment of the intended regulation types proposed by SDT; and (2) absence of conceptual redundancy with other items. Based on these five criteria, the group decided to delete 10 items. The shortened scale, which is presented in the Appendix A, includes 15 items (three per regulation type). On this scale, a general question first asks participants to rate the extent to which each item corresponds to their reasons for persisting in their doctoral studies on a five-point Likert scale (1 = does not correspond at all, 5 = corresponds exactly).

2.3. Measures

In addition to the MPhD scale, participants were asked to complete measures related to their dropout intentions (both samples), intentions to pursue postdoctoral research (both samples), satisfaction with their studies (both samples) and with their universities and programs more specifically (Sample 2 only), their levels of test anxiety and positive–negative affect (Sample 1 only), and their levels of academic performance and problems encountered in relation to their thesis (Sample 2 only).

2.3.1. Dropout intention (both samples)

Based on Schmitz et al. (2010), two items were used to assess students' dropout intention in both samples. We asked the following question: "What is the likelihood that you will give up your studies?" Participants answered on a five-point Likert scale (1 = *not at all likely*, 5 = *very likely*) for two possibilities; 1) "in the next few months" and 2) "prior to graduation." Cronbach's alpha for this scale was .77 in Sample 1 and .91 in Sample 2. As the scale includes only two items, the Spearman–Brown formula was used to obtain a corrected estimate of reliability. The adjusted scale score reliability coefficient for this scale estimated for eight equivalent items was .93 in Sample 1 and .98 in Sample 2. A distinction was made between these two possibilities, as doctoral students who quit their program do so after completing an average of nearly three years (MERS, 2013). The decision to quit appears to be a long process, and it is plausible that a student stays in a program for another year despite believing that he or she will not complete it (e.g., enjoying a prestigious scholarship while waiting for an interesting job opportunity).

2.3.2. Postdoctoral intention (both samples)

One dichotomous ("yes" or "no") item was used to assess postdoctoral intention ("After your doctorate, do you plan to pursue postdoctoral research?") in both samples. Nearly half the participants (46% in Sample 1, 45% in Sample 2) answered positively.

2.3.3. Satisfaction with studies (general satisfaction; both samples)

An adaptation of the French version of the Satisfaction with Life Scale (*Échelle de satisfaction de vie*; Blais, Vallerand, Pelletier, & Brière, 1989) was used to assess students' satisfaction with their studies in both samples. This instrument contains four items (e.g., "I am satisfied with my studies") rated on a seven-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). Cronbach's alpha for this scale was .86 in Sample 1 and .83 in Sample 2.

2.3.4. Satisfaction with the university (university satisfaction; Sample 2)

Four items were used, in Sample 2 only, to assess satisfaction with the university and the field of study (e.g., "If you had to start your PhD studies over, would you choose the same university?," "Would you recommend this university to someone interested in another field of study?"). Responses were rated on a five-point Likert scale (1 = *certainly not*, 5 = *without a doubt*). Cronbach's alpha was .74.

2.3.5. Satisfaction with the program (program satisfaction; Sample 2)

A total of 16 items were used, in Sample 2 only, to assess participants' satisfaction with their program. These items were inspired by the Canadian Graduate and Professional Student Survey, a national survey conducted by the Canadian Association of Graduate Studies (2010). Each item represents an aspect of the program, and participants were asked to rate them on a five-point Likert scale (1 = *poor*, 5 = *excellent*). Examples of items are, "Professors' competence level," "Availability of faculty members outside class hours," and "Relation between program content and my research objectives." Cronbach's alpha was .90.

2.3.6. Test anxiety (Sample 1)

We used an adaptation of the Test Anxiety Scale (*Échelle d'anxiété envers l'évaluation*; Beaudoin & Desrichard, 2009) to assess test anxiety, in Sample 1 only. This instrument contains six items (e.g., "When I think about my future assessment, I feel anxious"). Participants rated each item on a seven-point Likert scale (1 = *does not correspond at all*, 5 = *corresponds exactly*). Cronbach's alpha was .87.

2.3.7. Positive and negative affect (Sample 1)

We administrated the Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) to assess participants' affectivity, in Sample 1 only. This scale contains 20 emotion adjectives, of which 10 assess dispositional positive affect and 10 measure dispositional negative affect. Participants rated the extent to which they felt these emotions on a five-point scale (1 = *not at all*, 5 = *completely*). Cronbach's alpha was .89 for positive affect and .84 for negative affect.

2.3.8. Academic performance (performance; Sample 2)

In Sample 2 only, participants were asked to rate their performance in their PhD studies in terms of grades, thesis project and research productivity on a five-point Likert scale (1 = *poor*, 5 = *excellent*). Cronbach's alpha was .73.

2.3.9. Thesis problems (Sample 2)

In Sample 2 only, four questions were used to assess the problems that individuals encountered in completing their thesis. Examples of items are, "Choosing a thesis topic is or was a difficult task," and "I have or had problems developing my thesis project." Participants rated each item on a seven-point Likert scale (1 = *completely disagree*, 7 = *completely agree*). Cronbach's alpha was .65.

3. Statistical analyses

3.1. Exploratory structural equation modeling (ESEM)

To test the factor structure of the MPhD scale, we conducted exploratory structural equation modeling (ESEM) using Mplus 7.3 (Muthén & Muthén, 2012). This recent statistical technique incorporates features of both confirmatory factor analysis (CFA) and exploratory factor analysis (EFA). Like CFA, ESEM tests whether the scale comprises five distinct factors and provides fit indices, standard errors, and tests of significance. However, it is less restrictive than CFA, relaxing the restrictive assumption that items should load only on their respective factors (i.e., main loading) without any cross-loading (Marsh, Morin, Parker, & Kaur1, 2014, Marsh et al., 2009; Morin, Marsh, & Nagengast, 2013). A measurement instrument may have many cross-loadings (albeit much weaker than their main loadings) that are consistent with the underlying theory, as is the case for SDT. For instance, because all the regulation types reflect motivation, some positive cross-loadings are expected, especially among conceptually adjacent factors (Guay, Morin, Litalien, Valois, and Vallerand, 2014). Moreover, when true cross-loadings (i.e. present in the population model) are forced to be zero in CFA, latent factor correlations tend to be overestimated, as the only way for the cross-loadings to be expressed is through the inflation of these correlations (e.g., Asparouhov & Muthén, 2009; Marsh et al., 2009; Morin, Arens, & Marsh, in press; Morin et al., 2013).

By incorporating cross-loadings in a model, an ESEM approach overcomes these limitations. It also provides some control over the fact that items are imperfect indicators of a construct, and thus presents some degree of irrelevant association with other constructs (i.e., systematic measurement error; see Morin et al., in press). ESEM thus appeared particularly relevant for investigating the psychometric properties of the MPhD scale and for estimating purer correlations among the latent variables in order to better assess the simplex pattern suggested by SDT.

To test the higher-order structure with ESEM, we used the ESEM-within-CFA (EWC) approach described by Morin et al. (2013, in press) (it is currently not possible to apply ESEM methodology to analyze higher-order factor structure given that higher-order rotational procedures have yet to be developed). In this method, parameter estimates from the final first-order ESEM solution are used as start values in the EWC model estimation. The same number of constraints as the ESEM model is added for identification purposes. An EWC solution typically has the same parameter estimates, goodness-of-fit, and degrees-of-freedom as the corresponding ESEM solution, but allows higher-order factors to be estimated.

Despite choosing ESEM analysis, we first ran CFA to test the five-factor structure, as recommended by Marsh et al. (2009). If the analysis reveals adequate and similar fit indices for both ESEM and CFA models, there is less advantage to pursuing an ESEM analysis because the ESEM model is less parsimonious than the CFA model. Nevertheless, an ESEM model can still provide a more exact representation of the factor correlations when cross-loadings are present in the population model (for a review, see Morin et al., 2013).

3.2. Estimator, missing data, and goodness-of-fit indices

All analyses were performed using the Mplus 7.3 (Muthén & Muthén, 2012) robust weighted least square (WLSMV) estimator for ordered-categorical variables. The models were estimated based on the full information that was available using algorithms implemented in Mplus in conjunction with the WLSMV estimator (Asparouhov & Muthén, 2010). WLSMV was chosen because this estimator is more suited to the ordered-categorical nature of Likert scales than traditional maximum likelihood estimation (Beauducel & Herzberg, 2006; Finney & DiStefano, 2006; Lubke & Muthén, 2004), resulting in more accurate estimates of key model-parameters. To assess model fit, we used the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), and the chi-square test statistic. It should be noted that previous research has shown that traditional fit indices (CFI, TLI, and RMSEA) perform quite well when the WLSMV estimator is used (Beauducel & Herzberg, 2006; Yu, 2002). Following recommendations from Marsh et al. (2009, 2010) (also see Guay et al., 2014; Morin et al., 2013), we used an oblique Geomin rotation with an epsilon value of 0.5.

3.3. Scale score reliability

Scale score reliability estimates were computed from the standardized parameter estimates of the models, using McDonald's (1970) omega, $\omega = (\Sigma|\lambda i|)^2/([\Sigma|\lambda i|]^2 + \Sigma\delta ii)$, where λi are the standardized factor loadings and δii, the standardized item uniquenesses. Compared with traditional scale score reliability estimates (e.g., alpha; see Sijtsma, 2009), ω has the advantage of taking into account the strength of association between items and constructs (λi) as well as item-specific measurement errors (δii).

3.4. Multigroup analyses

Invariance of the measurement model across both samples and meaningful subgroups of PhD students (i.e., gender, citizenship status, program type, age, and study progression) was tested based on the sequence described in Guay et al. (2014) for ordered-categorical items. For each subgroup comparison, we successively assessed: (1) configural invariance, (2) weak (loadings) invariance, (3) strong (loadings, thresholds) invariance, (4) strict (loadings, thresholds, uniquenesses) invariance, (5) invariance of the variance/covariance matrix (loadings, thresholds, uniquenesses, latent variances and covariances), and (6) latent mean invariance (loadings, thresholds, uniquenesses, latent variances and covariances, latent means). Both chi-square difference tests conducted using the Mplus DIFFTEST function (MD$\Delta\chi2$; Asparouhov & Muthén, 2006) and fluctuations in fit indices were used to compare nested models. The invariance hypothesis should not be rejected for the nested model when the CFI decrease is .01 or less or when the RMSEA increase is .015 or less (Chen, 2007; Cheung & Rensvold, 2002).

4. Results

4.1. Sample 1

4.1.1. Preliminary verifications

Descriptive statistics for motivation items are presented in Table 1. Items were normally distributed, presenting skewness and kurtosis indices within the limits suggested by West, Finch, and Curran (1995). Variable means and standard deviations are provided at the bottom of Table 2. Means for intrinsic, integrated, identified and external regulations, test anxiety, satisfaction with studies, and positive affect were higher than the scale midpoint. However, means were lower than the scale midpoint for introjected regulation, negative affect, and dropout intention.

4.1.2. Factor validity

The ESEM analysis used to test the a priori five-factor structure of the MPhD scale showed excellent fit to the data (χ^2 [40] = 53.385, CFI = .99, TLI = .99, and RMSEA = .04) and fit indices that were systematically superior to those obtained with the comparative CFA model (χ^2 [80] = 218.656, CFI = .94, TLI = .93, and RMSEA = .08). All items loaded strongly on their respective factors (ranging from .33 to .98, M = .69, SD = .19), and cross-loadings were systematically weaker than the main loadings (−.20 to .37, $|M|$ = .08; SD = .08; see Table 1). Two exceptions to this general pattern were observed: (1) the third item of intrinsic regulation [i.e., For the pleasure I feel in accomplishing my study project (e.g., thesis)] loaded almost equally on its a priori construct (.33) and on the identified regulation factor (.37); and (2) although item 1 of external regulation (i.e., For the prestige associated with a PhD) loaded more strongly on its a priori construct (.42) than on other factors, it also showed significant cross-loadings with all regulation types (and one negative cross-loading with identified regulation).

The higher-order EWC model showed excellent fit to the data (χ^2 [44] = 56.179, CFI = .99, TLI = .99, and RMSEA = .03), with indices similar to those obtained in the first-order ESEM model. As expected, intrinsic, integrated, and identified regulations loaded significantly on autonomous motivation while introjected and external regulations loaded significantly on controlled motivation (see Table 1).

4.1.3. Scale score reliability

Scale score reliability estimates (ω) for the subscales were .79 for intrinsic regulation, .85 for integrated regulation, .69 for identified regulation, .73 for introjected regulation, and .85 for external regulation.[1]

[1] Cronbach's alphas (α) were also computed and led to similar results (.73 for intrinsic regulation, .81 for integrated regulation, .65 for identified regulation, .61 for introjected regulation, and .76 for external regulation).

Table 1
Sample 1: Descriptive statistics and factor loadings from ESEM and EWC solutions.

Indicators	Descriptive statistics				Factor loadings				
	M	*SD*	Skewness	Kurtosis	1	2	3	4	5
Firstorder (ESEM)									
Intrinsic									
Item 1	3.86	1.03	−0.72	−0.07	.980**	.053	–.028	–.004	.027
Item 2	3.90	0.94	−0.80	0.55	.630**	.073	.196**	–.037	–.069
Item 3	3.88	1.03	−0.77	0.12	.328**	.081	.367**	–.196**	–.108
Integrated									
Item 1	3.99	0.96	−0.81	0.18	.145**	.791**	–.072	–.045	.021
Item 2	3.43	1.21	−0.35	−0.79	.046	.803**	.090*	.012	–.010
Item 3	3.86	1.08	−0.86	0.16	.080	.698**	.186**	–.027	–.002
Identified									
Item 1	4.14	0.84	−1.11	1.89	.050	.123*	.559**	–.164**	.114*
Item 2	3.80	1.02	−0.63	−0.10	.043	.064	.746**	–.029	.004
Item 3	3.44	1.12	−0.36	−0.63	.191**	.045	.536**	.073	–.060
Introjected									
Item 1	1.55	0.87	1.54	1.61	–.040	–.064	.060	.737**	.008
Item 2	2.24	1.22	0.78	−0.34	–.101	–.079	.136	.480**	–.145*
Item 3	1.82	1.07	1.25	0.69	–.025	.009	–.156**	.792**	.054
External									
Item 1	2.62	1.30	0.44	−0.86	.201**	.191**	–.200**	.368**	.417**
Item 2	3.20	1.23	−0.30	−0.87	–.022	–.040	.031	–.046	.914**
Item 3	2.96	1.29	0.10	−1.10	–.013	.014	.012	.044	.897**
Higher–order (EWC)									
Autonomous motivation									
Intrinsic regulation	3.88	0.81	−0.77	0.49	.652**	–	–	–	–
Integrated regulation	3.76	0.93	−0.58	−0.20	.524**	–	–	–	–
Identified regulation	3.79	0.77	−0.57	0.06	.394**	–	–	–	–
Controlled motivation									
Introjected regulation	1.87	0.80	1.10	0.95	–	.973**	–	–	–
Controlled regulation	2.93	1.05	0.13	−0.76	–	.274*	–	–	–

Note. Grayscale indicates loadings of the items on their target a priori factor; ** = $p < .01$. * = $p < .05$.

4.1.4. Convergent and discriminant validity

First, correlations among the regulation subscales were assessed from the ESEM model (see Table 2, under the diagonal). They were generally lower than those estimated with CFA (see Table 2, above the diagonal). As expected, correlations (especially ESEM ones) were mostly in line with the simplex pattern postulated by SDT: adjacent regulation types on the continuum were positively associated, and distant regulation types were less strongly associated. Intrinsic regulation was positively associated with integrated ($r = .37$) and identified regulation ($r = .30$), negatively associated with introjected regulation ($r = -.11$),

Table 2
Sample 1: correlations, means, and standard deviations for all variables.

Factor	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13
Five-factor structure (ESEM)													
F1. Intrinsic regulation	–	.59**	.70**	−.36**	−0.07								
F2. Integrated regulation	.37**	–	.49**	−.21**	0.09								
F3. Identified regulation	.30**	.22**	–	−.36**	−0.10								
F4. Introjected regulation	−.11*	−0.07	−.22**	–	.23**								
F5. External regulation	0.01	0.08	−.11*	.11*	–								
Higher-order structure (EWC)													
F6. Autonomous motivation	–	–	–	–	–	–							
F7. Controlled motivation	–	–	–	–	–	−.02	–						
Outcomes													
F8. Test anxiety	−.13*	−.14*	−.17*	.26**	.13*	−.25**	.31**	–					
F9. General satisfaction	.41**	.39**	.41**	−.31**	−.21**	.70**	−.39**	−.22**	–				
F10. Positive affect	.54**	.47**	.55**	−.39**	−.03	.53**	−29**	−.42**	.61**	–			
F11. Negative affect	−.13*	−.16*	−.17*	.42**	.11	−.28**	.79**	.63**	−.37**	−.43**	–		
F12. Dropout intention	−.25**	−.25**	−.15	.24**	−.02	−.39**	0.18	.26*	−.41**	−.45**	.23**	–	
F13. Postdoctoral intention	.12	.17*	.31**	−.18	−.11	.33*	−0.23	−.05	.19*	.23*	−.20*	−.16	–
Descriptive statistics													
Mean (*M*)	3.88	3.76	3.79	1.87	2.93	3.81	2.4	3.67	5.05	3.64	2.10	1.52	.46
Standard deviation (*SD*)	.81	.93	.77	.80	1.05	0.67	0.7	.96	1.22	.71	.71	.62	.50

Note. Correlations below the diagonal were obtained from ESEM and EWC solutions. Correlations above the diagonal were obtained from a CFA solution. Descriptive statistics were obtained using SPSS. Postdoctoral intention was assessed by one dichotomous item (no = 0; yes = 1). ** = $p < .01$. * = $p < .05$.

and not associated with external regulation. Integrated regulation was also positively associated with identified regulation ($r = .22$) and not associated with introjected and external regulations. Identified regulation was negatively associated with both introjected ($r = -.22$; unexpected result) and external ($r = -.11$) regulations, while introjected regulation was positively associated ($r = .11$) with external regulation.

Second, the correlation between the higher-order factors was estimated with the EWC model. The results showed that autonomous and controlled motivations were not associated (i.e., orthogonal; $r = -.02$). Third, correlations among the regulation types and the various outcomes considered here were obtained by adding covariates, defined as CFA factors, to the final retained models. Relationships between the covariates and the first-order regulation subscales were tested using the first-order ESEM model. Then, relationships between the covariates and the broader motivation categories (i.e., the higher-order factors) were assessed using the higher-order model (EWC). Both of these models provided adequate fit to the data (CFI and TLI ≥ .90, RMSEA ≥ 0.06) and correlation estimates mostly in line with the hypotheses (see Table 2).

Third, correlations among the regulation types and the various outcomes considered here were obtained by adding covariates, defined as CFA factors, to the final retained models. Relationships between the covariates and the first-order regulation subscales were tested using the first-order ESEM model. Then, relationships between the covariates and the broader motivation categories (i.e., the higher-order factors) were assessed using the higher-order model (EWC). Both of these models provided adequate fit to the data (CFI and TLI ≥ .90, RMSEA ≥ 0.06) and correlation estimates mostly in line with the hypotheses (see Table 2).

At the first-order level, autonomous types of regulation (intrinsic, integrated, identified) were positively associated with satisfaction with studies and positive affect, and were negatively associated with test anxiety and negative affect. In addition, intrinsic and integrated regulations were negatively associated with dropout intention, whereas integrated and identified regulations were positively associated with postdoctoral intention. Correlations were not stronger for intrinsic regulation than for integrated or identified regulations. Concerning controlled types of regulation, introjected regulation was negatively associated with satisfaction with studies and positive affect, and positively associated with test anxiety, negative affect, and dropout intention. External regulation was negatively associated with satisfaction with studies and positively associated with test anxiety, but no association was found with other outcomes.

At the higher-order level, autonomous motivation was positively associated with satisfaction with studies, positive affect, and postdoctoral intention, but negatively associated with text anxiety, negative affect, and dropout intention. Controlled motivation showed an opposite pattern for text anxiety, satisfaction with studies, positive and negative affect, but was unrelated to dropout or postdoctoral intentions.

4.2. Sample 2

4.2.1. Preliminary verifications

Descriptive statistics for the motivation scale items are reported in Table 3. Items were normally distributed, presenting skewness and kurtosis indices within the limits suggested by West et al. (1995). Percentages of missing data on variables ranged from 0% to 15% (see Table 4). Factor means and standard deviations are also presented at the bottom of Table 4. Means for intrinsic, integrated, identified and external regulations, test anxiety, satisfaction with studies, and positive affect were higher than the scale midpoint. Means were lower than the scale midpoint for introjected regulation, negative affect, and intention to dropout.

4.2.2. Factor validity

Once again, the a priori ESEM representation showed excellent fit to the data: (χ^2 [40] = 123.427, CFI = .99, TLI = .98, and RMSEA = .04), and fit indices that were superior to those obtained with CFA (χ^2 [80] = 952.568, CFI = .91, TLI = .88, and RMSEA = .10). Most items loaded strongly on their a priori factors (ranging from .23 to .93, $M = .63$, $SD = .19$) and cross-loadings were systematically weaker than the main loadings (−.32 to .53, $|M| = .09$; $SD = .10$; see Table 3). Two exceptions to this general pattern were observed, and involved the same items as in Sample 1: (1) the third item of intrinsic regulation loaded at .34 on its construct and at .53 on identified regulation, and (2) the first item of external regulation loaded at .23 on its construct, but also slightly higher on intrinsic (.28), integrated (.30), and introjected (.24) regulations. Overall, these results confirm a five-factor structure and suggest that two items should be targeted for reassessment and perhaps reformulation or replacement.

The higher-order EWC model showed excellent fit to the data (χ^2 [44] = 111.483, CFI = .99, TLI = .99, and RMSEA = .03), with indices similar to those obtained in the first-order ESEM model. As in Sample 1, intrinsic, integrated, and identified regulations loaded significantly on autonomous motivation, and introjected and external regulations loaded significantly on controlled motivation (see Table 3).

4.2.3. Scale score reliability

Scale score reliability estimates (ω) for the subscales were .73 for intrinsic regulation, .73 for integrated regulation, .60 for identified regulation, .78 for introjected regulation, and .81 for external regulation.[2]

4.2.4. Convergent and discriminant validity

Correlations among the regulation subscales partly corroborated the expected simplex pattern (see ESEM analysis results, Table 4, under the diagonal), but not as clearly as in Sample 1. Intrinsic regulation was positively associated with all other regulation types, and the strength of the correlations followed the expected SDT continuum (integrated, $r = .45$; identified, $r = .27$; introjected, $r = .15$; external, $r = .12$). Integrated regulation was also associated with identified ($r = .28$) and external ($r = .20$) regulations following a similar pattern. Identified regulation showed an unexpected pattern of correlations, as it was negatively associated with introjected regulation ($r = -.10$) but positively associated with external regulation ($r = .06$). Introjected and external regulation were positively associated ($r = .20$). Unexpectedly, external regulation was positively, albeit weakly, associated with all autonomous types of regulation. It was also associated more strongly with intrinsic regulation ($r = .12$) than with identified regulation ($r = .06$). In the higher-order model, autonomous and controlled motivations were positively associated ($r = .31$).

As in Sample 1, correlations among the regulation types and the various outcomes considered here were obtained by adding covariates to the final models (first-order ESEM and higher-order EWC; see Table 4). Once again, both of these models provided adequate fit to the data (CFI and TLI ≥ .90, RMSEA ≤ 0.06) and correlation estimates mostly in line with the hypotheses.

At the first-order level, autonomous types of regulations (intrinsic, integrated, identified) were positively associated with general satisfaction, university satisfaction, program satisfaction, performance, and postdoctoral intention, and negatively associated with thesis problems. Integrated and identified regulations were also negatively associated with dropout intention. Introjected regulation was negatively associated with general satisfaction, university satisfaction, and performance, but positively associated with dropout intention and thesis problems. Finally, external regulation was negatively associated with performance and positively associated with thesis problems.

At the higher-order level, autonomous motivation was positively associated with general satisfaction, university satisfaction, program satisfaction, performance, and postdoctoral intention, and negatively

[2] Cronbach's alphas (α) were of a similar magnitude (.69 for intrinsic regulation, .71 for integrated regulation, .60 for identified regulation, .68 for introjected regulation, and .71 for external regulation).

Table 3
Sample 2: Descriptive statistics and factor loadings from ESEM and EWC solutions.

	Descriptive statistics				Factor loadings				
Item	*M*	*SD*	Skewness	Kurtosis	1	2	3	4	5
First–order (ESEM)									
Intrinsic									
Item 1	3.67	1.16	−0.71	−0.26	.755**	.145**	−.070**	−.017	.038*
Item 2	3.72	1.13	−0.80	−0.11	.711**	−.001	.198**	.051*	−.002
Item 3	3.89	1.08	−0.90	0.23	.341**	.053	.527**	−.048	−.087**
Integrated									
Item 1	4.24	0.91	−1.28	1.52	.213**	.553**	.023	−.057	−.011
Item 2	3.22	1.28	−0.26	−0.98	.145**	.659**	.055*	.097**	−.028
Item 3	3.77	1.15	−0.79	−0.15	.028	.667**	.194**	−.053*	.120**
Identified									
Item 1	4.17	0.96	−1.18	1.02	.069	.193**	.532**	−.076*	.231**
Item 2	3.91	1.06	−0.86	0.20	.031	.234**	.604**	−.019	−.012
Item 3	3.25	1.30	−0.39	−0.94	.157**	.024	.432**	.173**	.104**
Introjected									
Item 1	1.61	0.96	1.61	1.83	−.017	.012	−.002	.784**	.014
Item 2	2.37	1.33	0.53	−0.95	.024	−.068*	.091	.662**	.058*
Item 3	1.94	1.21	1.07	−0.01	.067	.053	−.126**	.719**	.081**
External									
Item 1	2.54	1.26	0.36	−0.94	.280**	.300**	−.326**	.244**	.228**
Item 2	3.23	1.31	−0.35	−1.00	.006	.036	−.013	−.017	.928**
Item 3	2.99	1.29	−0.11	−1.09	.014	−.014	.039*	.089**	.847**
Higher–order (EWC)									
Autonomous Motivation									
Intrinsic regulation	3.76	0.89	−0.73	0.15	.685**	–	–	–	–
Integrated regulation	3.74	0.90	−0.62	−0.12	.572**	–	–	–	–
Identified regulation	3.77	0.83	−0.73	0.40	.308**	–	–	–	–
Controlled motivation									
Introjected regulation	1.97	0.92	0.91	0.31	–	.293**	–	–	–
Controlled regulation	2.92	1.03	−0.07	−0.75	–	.664**	–	–	–

Note. Grayscale indicates loadings of the items on their target a priori factor; ** = $p < .01$. * = $p < .05$.

associated with dropout intention and thesis problems. Controlled motivation was negatively associated with general satisfaction and performance, positively associated with dropout intention and thesis problems, and unrelated to university and program satisfaction and postdoctoral intention.

4.3. Combined sample

4.3.1. Measurement invariance across samples 1 and 2

To verify the extent to which the final retained ESEM model was replicated across samples (and whether observed differences were

Table 4
Sample 2: Correlations, Means, standard deviations, and percentage of missing values for all variables.

Factor	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14
Five-factor structure (ESEM)														
F1. Intrinsic regulation	–	.72**	.74**	.13**	.23**									
F2. Integrated regulation	.45**	–	.70**	.09*	.34**									
F3. Identified regulation	.27**	.28**	–	0.03	.29**									
F4. Introjected regulation	15**	.05	−.10**	–	.40**									
F5. External regulation	.12**	.20**	.06**	.20**	–									
Higher-order structure (EWC)														
F6. Autonomous motivation	–	–	–	–	–	–								
F7. Controlled motivation	–	–	–	–	–	.31**	–							
Outcomes														
F8. General satisfaction	.27**	.36**	.54**	−.16**	.01	.72**	−15*	–						
F9. University satisfaction	.12**	.16**	.37**	−.10*	−.01	.41**	−12	.63**	–					
F10. Program satisfaction	.15**	.17**	.33**	−.05	−.06	.40**	−0.08	.58**	.66**	–				
F11. Performance	.24**	.22**	.33**	−.21**	−.08*	.47**	−.27**	.57**	.29**	.35**	–			
F12. Dropout intention	−.04	−.25**	−.24**	.23**	−.06	−.34**	.23**	−.47**	−.32**	−.27**	−.43**	–		
F13. Thesis problems	−.25**	−.20**	−.57**	.31**	.12**	−.63**	.43**	−.70**	−.49**	−.41**	−.69**	.50**	–	
F14. Postdoctoral intention	.11*	.34**	.33**	.00	.05	.47**	0.07	.14**	.01	.03	.15**	−.23**	−.16*	–
Descriptive statistics														
Mean (*M*)	3.76	3.74	3.77	1.97	2.92	3.76	2.45	5.09	3.95	3.48	3.67	1.57	2.86	0.46
Standard deviation (*SD*)	0.89	0.90	0.83	0.92	1.03	0.7	0.79	1.13	0.75	0.61	0.69	0.72	0.97	0.50
Missing values (%)	0	0	0	0	0	0	0	0	4.1	4.1	4.9	10.8	15.1	15.2

Note. Correlations below the diagonal were obtained from ESEM and EWC solutions. Correlations above the diagonal were obtained from a CFA solution. Descriptive statistics were obtained using SPSS. Postdoctoral intention was assessed by one dichotomous item (no = 0; yes = 1). Descriptive statistics were obtained using SPSS. ** = $p < .01$. * = $p < .05$.

due to random sampling errors), we systematically assessed the measurement invariance of the final ESEM and EWC models across Samples 1 ($N = 244$) and 2 ($N = 1060$). More restrictive models showed a good fit to the data (CFI and TLI ≥ .97, RMSEA ≤ 0.05; see M1 to M6 in Tables 5 and 6). The decrease in fit indices remained low and under recommended cut-off scores for most models, providing support for the strict measurement invariance of the model across samples (i.e., loadings, thresholds, and uniquenesses). The results also supported the invariance of the estimated latent means across samples. However, the results failed to support the invariance of the latent variances and covariances across samples (M5 in both Tables 5 and 6). A detailed examination of these results showed that four of the five motivation subscales and both higher-order factors showed a higher level of variability in Sample 2, which is consistent with the far larger size of this sample. Covariances did not differ across samples. Overall, these results confirm that the estimated measurement model (i.e. loadings, thresholds, and uniquenesess) was fully replicated across samples and that any observable differences in the parameter estimates associated with both models (see Tables 1 and 3) are likely due to random sampling error.

4.3.2. Measurement invariance across meaningful subgroups

Since the measurement model was found to be fully invariant across samples, and in order to maximize the available sample size for meaningful group comparisons, measurement invariance tests between meaningful subgroups of PhD students were conducted on the combined sample. Students were grouped by gender (males, $n = 522$; females, $n = 628$), citizenship status (Canadian, $n = 782$; permanent resident or temporary visa holder, $n = 368$), program type (social sciences, $n = 631$; natural sciences, $n = 652$), age (30 years old or less, $n = 796$; more than 30 years old, $n = 508$), and study progression (number of completed trimesters; 0 to 3, $n = 375$; 4 to 6, $n = 253$; 7 to 11, $n = 272$; 12 or more, $n = 249$). Measurement invariance of the first-order ESEM and higher-order EWC models was tested across all of these subgroups of students. All models showed a good fit to the data (CFI and TLI > .97, RMSEA < 0.05; see Tables 5 and 6). Furthermore, across all of the models estimated, as invariance constraints were added to the model, the observed decrease in fit indices remained under the recommended cut-off scores. These results thus fully support the strict measurement invariance of the MPhD scale across these subgroups of participants, as well as the invariance of the latent variances, covariances, and means of the estimated factors.

Table 5
Summary of goodness-of-fit statistics for the five-factor structure (ESEM) invariance tests.

Tested models	χ^2	df	*CFI*	*TLI*	*RMSEA* [90% *CI*]	*MD*$\Delta\chi^2$ (*df*)	Δ*CFI*	Δ*TLI*	Δ*RMSEA*
Sample									
M1. Configural	192.903*	80	.990	.974	.047 [.038, .055]	–	–	–	–
M2. Weak	200.832*	130	.994	.990	.029 [.021, .037]	63.228 (50)	+.004	+.016	−.018
M3. Strong	333.776*	170	.986	.983	.038 [.032, .045]	140.209 (40)*	−.008	−.007	+.009
M4. Strict	390.785*	185	.982	.980	.041 [.036, .047]	66.164 (15)*	−.004	−.003	+.003
M5. Var − Cov	533.282*	200	.971	.970	.051 [.045, .056]	94.391 (15)*	−.011	−.010	+.010
M6. Means#	349.416*	190	.986	.985	.036 [.030, .042]	5.699 (5)	+.015	+.015	−.015
Gender									
M7. Configural	170.537*	80	.991	.976	.044 [.035, .054]	–	–	–	–
M8. Weak	183.442*	130	.995	.991	.027 [.017, .035]	55.565 (50)	+.004	+.015	−.017
M9. Strong	284.422*	170	.988	.986	.034 [.027, .041]	103.838 (40)*	−.007	−.005	+.007
M10. Strict	305.341*	185	.988	.986	.034 [.027, .040]	24.664 (15)	.000	.000	.000
M11. Var–Cov	275.149*	200	.992	.992	.026 [.018, .033]	19.107 (15)	+.004	+.006	−.008
M12. Means	341.081*	205	.986	.986	.034 [.028, .040]	28.630 (5)*	−.006	−.006	.008
Citizenship status									
M13. Configural	163.823*	80	.992	.978	.043 [.033, .052]	–	–	–	–
M14. Weak	198.739*	130	.993	.989	.030 [.022, .039]	68.278 (50)*	+.001	+.011	−.013
M15. Strong	283.616*	170	.989	.986	.034 [.027, .041]	89.803 (40)*	−.004	−.003	+.004
M16. Strict	323.660*	185	.986	.984	.036 [.029, .043]	43.235 (15)*	−.003	−.002	+.002
M17. Var–Cov	339.977*	200	.986	.985	.035 [.028, .041]	38.725 (15)*	.000	+.001	−.001
M18. Means	408.353*	205	.980	.979	.042 [.036, .047]	32.940 (5)*	−.006	−.006	+.007
Program type									
M19. Configural	172.620*	80	.992	.978	.042 [.034, .051]	–	–	–	–
M20. Weak	226.570*	130	.991	.986	.034 [.027, .041]	85.031 (50)*	−.001	+.008	−.008
M21. Strong	319.881*	170	.987	.983	.037 [.031, .043]	98.922 (40)*	−.004	−.003	+.003
M22. Strict	339.792*	185	.986	.984	.036 [.030, .042]	25.133 (15)*	−.001	+.001	−.001
M23. Var–Cov	314.865*	200	.990	.989	.030 [.023, .036]	26.305 (15)*	+.004	+.005	−.006
M24. Means	377.561*	205	.985	.984	.036 [.030, .042]	28.988 (5)*	−.005	−.005	+.006
Age									
M25. Configural	193.749*	80	.990	.974	.047 [.038, .055]	–	–	–	–
M26. Weak	237.166*	130	.991	.985	036 [.028, .043]	85.281 (50)*	+.001	+.011	−.011
M27. Strong	362.532*	170	.983	.979	.042 [.036, .048]	134.691 (40)*	−.008	−.006	+.006
M28. Strict	463.891*	185	.976	.973	.048 [.043, .054]	97.967 (15)*	−.007	−.006	+.006
M29. Var–Cov	359.961*	200	.986	.985	.035 [.029, .041]	18.067 (15)	+.010	+.012	−.013
M30. Means	353.488*	205	.987	.987	.033 [.027, .039]	8.006 (5)	+.001	+.002	−.002
PhD progression									
M31. Configural	234.812*	160	.993	.981	.040 [.029, .051]	–	–	–	–
M32. Weak	387.678*	310	.992	.990	.030 [.019, .038]	184.890 (150)*	−.001	+.009	−.010
M33. Strong	574.290*	430	.986	.986	.034 [.026, .041]	201.371 (120)*	−.006	−.004	+.004
M34. Strict	659.107*	475	.982	.984	.037 [.030, .043]	92.192 (45)*	−.004	−.002	+.003
M35. Var–Cov	717.584*	520	.980	.984	.036 [.030, .043]	82.245 (45)*	−.002	.000	−.001
M36. Means	789.910*	535	.975	.980	.041 [.035, .047]	45.252 (15)*	−.005	−.004	+.005

Note. * = $p < .05$. # = The fit of the mean invariance model for sample invariance was compared to the fit of the strict invariance model, as the invariance of the latent variance–covariance matrix was not supported. Var–Cov = Variance–Covariance.

Table 6
Summary of goodness-of-fit statistics for the higher-order structure (EWC) invariance tests.

Tested models	χ^2	df	CFI	TLI	RMSEA [90% CI]	MD$\Delta\chi^2$ (df)	ΔCFI	ΔTLI	ΔRMSEA
Sample									
M1. Configural	433.820*	193	.979	.977	.044 [.038, .049]	–	–	–	–
M2. Weak	429.801*	196	.980	.978	.043 [.037, .048]	6.526 (3)	+.001	+.001	−.001
M3. Strong	418.783*	199	.981	.980	.041 [.036, .047]	2.486 (3)	+.001	+.002	−.002
M4. Strict	412.361*	204	.982	.982	.040 [.034, .045]	10.860 (5)	+.001	+.002	−.001
M5. Var–Cov	543.555*	207	.971	.971	.050 [.045, .055]	43.832 (3)*	−.011	−.011	+.010
M6. Means#	387.284*	206	.984	.984	.037 [.031, .042]	1.914 (2)	+.013	+.013	−.013
Gender									
M7. Configural	319.482*	193	.987	.986	.034 [.027, .040]	–	–	–	–
M8. Weak	313.174*	196	.988	.987	.032 [.025, .039]	3.428 (3)	+.001	+.001	−.002
M9. Strong	311.157*	199	.989	.988	.031 [.024, .038]	3.262 (3)	+.001	+.001	−.001
M10. Strict	313.527*	204	.989	.989	. 031 [.024, .037]	8.284 (5)	.000	+.001	.000
M11. Var–Cov	296.050*	207	.991	.991	.027 [.020, .037]	5.598 (3)	+.002	+.002	−.004
M12. Means	352.856*	209	.985	.985	.035 [.028, .041]	21.993 (5)*	−.006	−.006	+.008
Citizenship status									
M13. Configural	439.383*	193	.975	.973	.047 [.041, .053]	–	–	–	–
M14. Weak	415.034*	196	.978	.977	.044 [.038, .050]	0.978 (3)	+.003	+.004	−.003
M15. Strong	443.408*	199	.976	.974	.046 [.040, .052]	21.854 (3)*	−.002	−.003	+.002
M16. Strict	443.659*	204	.976	.975	.045 [.039, .051]	17.104 (5)*	.000	+.001	−.001
M17. Var-Cov	383.964*	207	.982	.982	.039 [.033, .045]	2.021 (3)	+.006	+.007	−.006
M18. Means	417.622*	209	.979	.979	.042 [.036, .047]	14.961 (2)*	−.003	−.003	+.003
Program type									
M19. Configural	354.937*	193	.986	.984	.036 [.030, .042]	–	–	–	–
M20. Weak	354.193*	196	.986	.985	.035 [.029, .041]	7.517 (3)	.000	+.001	−.001
M21. Strong	366.962*	199	.985	.984	.036 [.030, .042]	9.115 (3)*	−.001	−.001	+.001
M22. Strict	428.888*	204	.980	.979	.041 [.036, .047]	35.668 (5)*	−.005	−.005	+.005
M23. Var–Cov	381.936*	207	.984	.984	.036 [.031, .042]	5.424 (3)	+.004	+.005	−.005
M24. Means	389.167*	209	.984	.984	.037 [.031, .042]	7.719 (2)*	.000	.000	.001
Age									
M25. Configural	464.722*	193	.977	.974	.046 [.041, .052]	–	–	–	–
M26. Weak	435.761*	196	.979	.978	.043 [.038, .049]	3.722 (3)	+.002	+.004	−.003
M27. Strong	442.635*	199	.979	.978	.043 [.038, .049]	10.600 (3)*	.000	.000	.000
M28. Strict	414.760*	204	.982	.981	.040 [.034, .045]	3.093 (5)	+.003	+.003	−.003
M29. Var–Cov	372.886*	207	.986	.985	.035 [.029, .041]	2.918 (3)	+.004	+.004	−.005
M30. Means	359.140*	209	.987	.987	.033 [.027, .039]	2.032 (2)	+.001	+.002	−.002
PhD progression									
M31. Configural	680.347*	491	.981	.984	.037 [.030, .043]	–	–	–	–
M32. Weak	701.502*	500	.980	.983	.037 [.031, .044]	19.957 (9)*	−.001	−.001	.000
M33. Strong	736.355*	509	.977	.981	.039 [.033, .046]	25.746 (9)*	−.003	−.002	+.002
M34. Strict	772.614*	524	.975	.980	.041 [.034, .047]	37.487 (15)*	−.002	−.001	+.002
M35. Var–Cov	749.039*	533	.979	.983	.038 [.031, .044]	14.727 (9)	+.004	+.003	−.003
M36. Means	795.861*	539	.975	.980	.041 [.035, .047]	23.141 (6)*	−.004	−.003	+.003

Note. * = $p < .05$. # = The fit of the mean invariance model for sample invariance was compared to the fit of the strict invariance model, as the invariance of the latent variance–covariance matrix was not supported. Var–Cov = Variance–Covariance.

5. Discussion

The goal of this study was to develop and validate an SDT-based scale to assess motivation for PhD studies. Overall, the results provide good support for the psychometric properties of the MPhD scale and are consistent across samples. In line with SDT, both a five-factor first-order structure (intrinsic, integrated, identified, introjected, and external regulations) and a two-factor higher-order structure (autonomous and controlled motivations) were supported. Concerning the five-factor structure, ESEM analyses showed excellent fit to the data and loadings in line with theoretical expectations.

Only two items presented an unexpected loading and cross-loading pattern, which was similar across both samples. First, item 3 of the intrinsic regulation subscale [For the pleasure I feel in accomplishing my study project (e.g., thesis)] presented a high cross-loading on the identified regulation factor in addition to its own factor. Second, item 1 of the external regulation subscale (For the prestige associated with a PhD) presented high cross-loadings on most of the other factors and a lower main loading in Sample 2. Looking back at item 3 from the intrinsic regulation subscale, this pattern of results may be explained by the fact that it taps not only into "pleasure," which defines intrinsic regulation, but also into "accomplishment," which is related to identified regulation. Thus, although this cross-loading can be explain, it also suggests that this item should be targeted for re-assessment, and possible replacement (e.g., "For the pleasure of doing research," "Because I really enjoy my field of study") in future research. The second item, characterized by a weak main loading and multiple cross-loadings, appears more problematic. This could be related to the incorporation of "prestige" in this item. Indeed, whereas prestige represents some form of reward, it is located mainly within the student's self-image, over and above external contingencies. Therefore, this item should also be targeted for re-assessment and replacement. Possible replacement items could include other forms of external contingencies not currently covered in the MPhD scale (e.g., "Because I am afraid of my advisor's reaction if I quit," "Because I will lose some privileges (salary, fellowship, etc.)").

The results further showed that the scale score reliability estimates for the various subscales were generally satisfactory (ω = .69 to .85), with only the reliability of the identified regulation subscale falling below .70 in Samples 1 (ω = .69) and 2 (ω = .60). This estimate was lower than expected, especially in Sample 2, suggesting that this subscale should be examined more thoroughly in future research and used with caution in the context of research relying on fully latent variable models (e.g., Bollen, 1989) that provide control for measurement error. However, it should be kept in mind that this estimate is based on a three-item scale, and that scale score reliability is affected notably by the number of items present in a scale (Sijtsma, 2009; Streiner, 2003) so that it is often useful to provide adjustment of reliability estimates. In this study, scale score reliability estimates adjusted for eight equivalent items using the Spearman–Brown prophecy formula varied between 0.80 and 0.94 across studies and subscales. Moreover, the values obtained for identified regulation prior to this adjustment are similar to those observed with other well-established and well-validated motivation scales, such as the AMS (Vallerand et al., 1989, 1992, 1993, 1997) and the Self-Regulation Questionnaire (SRQ; Ryan & Connell, 1989).

Convergent and discriminant validity were assessed in two ways. First, correlations between the regulation subscales were mainly in line with the simplex pattern proposed by SDT. Some exceptions deserve our attention. Identified regulation was negatively associated with introjected regulation, whereas SDT posits that this correlation should be positive, although not very high. Other studies also support this hypothesis, showing a positive and significant association between these regulation types (Gagné et al., 2010; Ryan & Connell, 1989; Vallerand et al., 1989, 1992). However, the negative correlation observed is not very surprising. As mentioned earlier, SDT distinguishes between autonomous motivation, which leads to positive outcomes, and controlled motivation, which is associated with negative outcomes. Thus, on the self-determination continuum, the delimitation between these two broader types of regulation appeared between identified and introjected regulation. This "breakdown" of the continuum can be logically expressed as a negative correlation between identified and introjected regulation. Additionally, in Sample 2 only, external regulation was positively associated with all autonomous types of regulation, whereas SDT posits that these variables should be negatively associated. These associations are also reflected at the higher-order level, as a positive correlation is observed between autonomous and controlled motivations (Sample 2). These results might be related to the high cross-loadings observed with item 1 of external regulation and thus suggested that further studies should pay attention to the external regulation subscale.

Second, convergent and discriminant validity was assessed through correlations between each regulation type and various outcomes. The findings are particularly interesting, as they support the notion that different forms of motivation yield different consequences (Deci & Ryan, 2000). Results from both samples showed that intrinsic, integrated, and identified regulations were all beneficial, as 36 of 39 correlations were significant and in the expected direction. Indeed, they were positively associated with the positive outcomes (general satisfaction with studies, satisfaction with the university and with the program, positive affect, performance, and postdoctoral intention) and negatively associated with the negative outcomes (test anxiety, negative affect, dropout intention, and thesis problems). Conversely, introjected regulation seemed more detrimental, showing a reversed correlation pattern with the outcomes (10 of 13 correlations were significant). Specifically, introjected regulation was positively associated with all negative outcomes, and negatively associated with most positive outcomes (satisfaction with general studies and university, positive affect, and performance).

Although introjected regulation could be considered deleterious, this was less obvious for external regulation, as only 4 of 13 correlations were significant and in the expected direction. In Sample 1, external regulation was positively related to test anxiety and negatively associated with satisfaction with studies, whereas in Sample 2, it was positively associated with thesis problems and negatively associated with performance. Nevertheless, external regulation was unrelated to the other outcomes like positive and negative affect, university and program satisfaction, dropout intention, and postdoctoral intention.

For doctoral students, introjected regulation appeared to be more counterproductive than drive to attain prestige, money, and good working conditions. This milder effect of external regulation was also obtained by Gagné et al. (2010; Motivation at Work Scale) in a population of workers. It is also plausible that a wider selection of assessed outcomes or items could have captured the negative effect of external regulation more clearly. For instance, depth of learning and creativity are relevant components at the PhD level and have been negatively associated with external motives in previous studies with other populations (Amabile, 1985; Vansteenkiste et al., 2004).

In both samples, the goodness-of-fit of the higher-order solution was also excellent and did not differ significantly from that of the first-order solution. The relations between the five types of regulation could also be explained by the two broader categories of motivation, which is consistent with SDT. Moreover, correlations between the outcomes and autonomous motivation were all significant and in the expected direction, while 8 of 13 correlations were similarly significant (and also in the expected direction) for controlled motivation. These results suggest that the MPhD scale could be used to assess either richer (five-factor structure) or more parsimonious (higher-order structure) regulation types. To make an informed choice, we suggest always testing the five-factor structure before using the more parsimonious one. For instance, in the present study, the three autonomous types of regulation clearly showed similar patterns of association with the various outcomes, whereas the patterns of association of the two controlled motivation types were more distinct. Introjected regulation was more deleterious than the other regulation types, and external regulation did not appear as harmful as initially expected.

The population of PhD students is heterogeneous and the PhD experience may differ greatly depending on their gender, citizenship status, program type, age, and progression. Interestingly, measurement invariance for the MPhD scale was supported across both samples and across subgroups formed from those variables. These results support the generalizability of the five-factor structure and the higher-order structure and suggest that the scale can be used with PhD students from various backgrounds.

Overall, our findings underscore the fact that different types of regulation should be assessed separately (either as a five-factor structure or a higher-order structure), as they are linked to distinct consequences. Although studies on doctoral students' persistence often conceptualize motivation from a quantitative perspective (presence or absence), our results suggested instead that special attention should be paid to the quality of motivation.

5.1. Limitations and future studies

Although this study generally supported the psychometric properties of our scale, some limitations warrant attention. First, the measures were self-reported, which increases common method variance (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). It is important to note, however, that self-reports also offer a number of advantages, and may be the method of choice for exploring intrapsychic factors, such as motivation (Crockett, Schulenberg, & Petersen, 1987; Howard, 1994). Nonetheless, the reliability of the scale should be further validated through testing over time and other methods, such as reports from significant others (e.g., an advisor) or diaries. Second, because participation in our study was voluntary, it is plausible that the participants were the most motivated of the PhD students who were approached. Third, the MPhD scale was validated only among French–Canadian students. Thus, the psychometric properties of this instrument should be cross-validated with students from various academic, cultural and linguistic backgrounds. Fourth, although convergent and discriminant validity were assessed, no criterion measures were collected at a later time. Further studies should test the predictive validity of the scale to explain persistence, progression, dropout, publication, and even publication impact using longer-term longitudinal follow-up studies of student cohorts.

5.2. Scientific significance of the study

Despite these limitations, our findings have methodological, theoretical and practical implications. Methodologically, we have provided a new instrument to assess motivation in PhD students. Furthermore, our scale assesses integrated regulation, a more advanced psychological development that is rarely addressed in educational studies (Ratelle et al., 2007). This type of regulation can be differentiated and is relevant to the doctoral experience. Theoretically, our findings support SDT assumptions in PhD students. On the one hand, we distinguished five types of regulation in a first-order structure and two broader types of motivation in a higher-order one. On the other hand, even at the PhD level, students experience better outcomes when their motivation is autonomous rather than controlled. Assessing the quality of motivation through the lens of SDT could therefore help provide a deeper understanding of PhD students' motivation and persistence. In practical terms, our findings open the way to the development of potentially effective interventions, suggesting that increasing autonomous motivation (intrinsic, integrated, and identified) and reducing controlled motivation (mostly introjected, based on our results) among PhD students could enhance their psychological well-being and reduce dropout intentions. In line with SDT, providing a supportive context could facilitate the internalization of controlled types of regulation. For instance, dissertation advisors should be aware of these findings, especially with regard to introjected regulation. Whereas some may believe that putting a lot of pressure on their students is a beneficial strategy for their progression and experience, our results suggest otherwise. Similarly, our correlational results further suggest that all regulation types are associated with thesis difficulties, test anxiety, and performance. Although longitudinal studies would be needed to test those predictive links more precisely, it seems likely that providing additional services to help students deal with dissertation problems or test anxiety would positively affect their motivation. Moreover, the MPhD scale could also be used to assess the effectiveness of programs and interventions by monitoring possible fluctuations in regulation types.

Appendix A. The MPhD scale

Does not correspond at all	Corresponds somewhat	Corresponds moderately well	Corresponds well	Corresponds exactly
1	2	3	4	5

The following 15 statements correspond to reasons that can motivate doctoral students to persevere in their studies. We ask you to indicate the extent to which each statement corresponds to the reasons why you persevere in your doctoral studies.

Motivation type and item number	Statements
Intrinsic	
Item 1	For the satisfaction I feel when I surpass myself in my learning activities (e.g., work, presentations).
Item 2	For the satisfaction I have in facing challenges in my studies.
Item 3	For the pleasure I feel in accomplishing my study project (e.g., thesis).
Integrated	
Item 1	Because doctoral studies are consistent with my values (e.g., curiosity, ambition, success).
Item 2	Because my doctoral studies are a fundamental part of who I am and my identity.
Item 3	Because my doctoral studies meet my goals and my objectives in life.
Identified	
Item 1	Because I want to improve my skills in my field of study.
Item 2	Because it's important for me to advance knowledge in my field of study.
Item 3	Because I have the opportunity to take my first steps in research (e.g., publications, collaborations) while benefitting from supervision.
Introjected	
Item 1	Because my supervisor would be disappointed or angry if I gave up.
Item 2	Because I have made commitments that I must fulfill (e.g., with funding agencies, employers, collaborators, a research director).
Item 3	Because I do not want to be perceived as a quitter.
External	
Item 1	For the prestige associated with a PhD.
Item 2	To find a job with good working conditions.
Item 3	To get a better paying job after graduation.

References

Ahmed, W., & Bruinsma, M. (2006). A structural model of self-concept, autonomous motivation and academic performance in cross-cultural perspective. *Electronic Journal of Research in Educational Psychology*, *4*, 551–576.

Amabile, T. M. (1985). Motivation and creativity: Effects of motivational orientation on creative writers. *Journal of Personality and Social Psychology*, *48*, 393–399.

Anderson, M. S., & Swazey, J. P. (1998). Reflections on the graduate student experience: An overview. In M. S. Anderson (Ed.), *The experience of being in graduate school* (pp. 3–27). San Francisco, CA: Jossey-Bass Publishers.

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

Asparouhov, T., & Muthén, B. O. (2006). *Robust chi-square difference testing with mean and variance adjusted test statistics.* Los Angeles, CA: Muthén & Muthén (Retrieved from http://www.statmodel.com/examples/webnote.shtml#web10).

Asparouhov, T., & Muthén, B. O. (2010). *Weighted least squares estimation with missing data.* Los Angeles, CA: Muthén & Muthén (Retrieved from http://www.statmodel.com/download/GstrucMissingRevision.pdf).

Association of Universities and Colleges of Canada (2009). Building a competitive advantage for Canada (pre-budget submission to the Hon. James Flaherty, Minister of Finance). Retrieved from http://www.aucc.ca/media-room/publications/building-a-competitive-advantage-for-canada

Auriol, L. (2010). Careers of doctorate holders: Employment and mobility patterns. *OECD science, technology and industry working papers. No. 2010/04*. OECD Publishinghttp://dx.doi.org/10.1787/5kmh8phxvvf5-en.

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *Journal of Higher Education*, *73*, 94–122.

Bair, C., & Haworth, J. (2005). Doctoral student attrition and persistence: A meta-synthesis of research. *Higher Education: Handbook of Theory and Research*, *19*, 481–534.

Beaudoin, M., & Desrichard, O. (2009). Validation of a short French state test worry and emotionality scale. *Revue Internationale de Psychologie Sociale*, *22*, 79–105.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, *13*, 186–203.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*, 740–756.

Blais, M. R., Vallerand, R. J., Pelletier, L. G., & Brière, N. M. (1989). L'échelle de satisfaction de vie: Validation canadienne-française du "Satisfaction with Life Scale". [The satisfaction scale: French–Canadian validation of the Satisfaction with Life Scale]. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, *21*, 210–223.

Bloom, D. E., Hartley, M., & Rosovsky, H. (2006). Beyond private gain: The public benefits of higher education. In J. J. F. Forest, & P. G. Altbach (Eds.), *International handbook of higher education* (pp. 293–308). Dordrecht, Netherlands: Springer.

Boiché, J., Sarrazin, P., Grouzet, F. M. E., Pelletier, L. G., & Chanal, J. (2008). Students' motivational profiles and achievement outcomes in physical education: A self-determination perspective. *Journal of Educational Psychology*, *100*, 688–701.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York, NY: Wiley.

Broder, H. L., McGrath, C., & Cisneros, G. J. (2007). Questionnaire development: face validity and item impact testing of the Child Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology*, *35*, 8–19.

Canadian Association of Graduate Studies (2010). Canadian Graduate and Professional Student Survey — 2010. Instrument for students in doctoral stream/research programs. Retrieved from https://www.dropbox.com/s/rgobhayaqm8p3r6/EN_REG_2010.pdf

Cardona, J. J. (2013). Determined to succeed: Motivation towards doctoral degree completion (doctoral dissertation). Retrieved from http://deepblue.lib.umich.edu/

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233–255.

Clewell, B. C. (1987). Retention of Black and Hispanic doctoral students (ETS research report no. 87-10). Retrieved from ETS website http://www.ets.org/Media/Research/pdf/RR-87-10-Clewell.pdf

Crockett, L., Schulenberg, J., & Petersen, A. (1987). Congruence between objective and self-report data in a sample of adolescents. *Journal of Adolescent Research*, *2*, 383–392.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* New York, NY: Plenum Press.

Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Deci, E. L., & Ryan, R. M. (2012). Motivation, personality, and development within embedded social contexts: An overview of self-determination theory. In R. M. Ryan (Ed.), *The oxford handbook of human motivation* (pp. 85–107). New York, NY: Oxford University Press.

Deemer, E. D., Martens, M. P., & Buboltz, W. C. (2010). Toward a tripartite model of research motivation: Development and initial validation of the Research Motivation Scale. *Journal of Career Assessment*, *18*, 292–309.

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin*, *125*, 276–302.

Elliot, A. J. (1997). Integrating the "classic" and "contemporary" approaches to achievement motivation: A hierarchical model of approach and avoidance achievement motivation. In M. L. Maehr, & P. R. Pintrich (Eds.), *Advances in motivation and achievement*, *Vol. 10.* (pp. 143–179). Greenwich, CT: JAI Press.

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 269–314). Greenwich, CT: Information Age.

Gagné, M., Forest, J., Gilbert, M. -H., Aubé, C., Morin, E., & Malorni, A. (2010). The Motivation at Work Scale: Validation evidence in two languages. *Educational and Psychological Measurement*, *70*, 628–646.

Gillet, N., Lafrenière, M. -A. K., Vallerand, R. J., Huart, I., & Fouquereau, E. (2014). The effects of autonomous and controlled regulation of performance-approach goals on well-being: A process model. *British Journal of Social Psychology*, *53*, 154–174.

Guay, F., Morin, A. J. S., Litalien, D., Valois, P., & Vallerand, R. J. (2014). Application of exploratory structural equation modeling to evaluate the Academic Motivation Scale. *The Journal of Experimental Education*, *83*, 51–82.

Guay, F., Ratelle, C. F., & Chanal, J. (2008). Optimal learning in optimal contexts: The role of self-determination in education. *Canadian Psychology/Psychologie Canadienne*, *49*, 233–240.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*, 47–77.

Hoskins, C. M., & Goldberg, A. D. (2005). Doctoral student persistence in counselor education programs: Student-program match. *Counselor Education and Supervision*, *44*, 175–188.

Howard, G. (1994). Why do people say nasty things about self-reports? *Journal of Organizational Behavior*, *15*, 399–404.

Human Resources and Skills Development Canada (2006). Looking ahead: A ten-year outlook for the Canadian labour market, (2006–2015) (catalogue no. HS28-23/2006E). Retrieved from http://www.hrsdc.gc.ca/eng/publications_resources/research/categories/labour_market_e/sp_615_10_06/page00.shtml

Ivankova, N., & Stick, S. (2007). Students' persistence in a distributed doctoral program in educational leadership in higher education: A mixed methods study. *Research in Higher Education*, *48*, 93–135.

Jablonski, A. M. (2001). Doctoral studies as professional development of educators in the United States. *European Journal of Teacher Education*, *24*, 215–221.

Kärner, A., Kukemelk, H., & Herdlein, R. J. (2005). Motivation for obtaining the doctor of philosophy degree in the post-soviet era: The case of Estonia. *International Education*, *35*, 24–34.

King, S. E., & Chepyator-Thomson, J. R. (1996). Factors affecting the enrollment and persistence of African–American doctoral students. *Physical Educator*, *53*, 170–180.

Litalien, D., Lüdtke, O., Parker, P., & Trautwein, U. (2013). Different pathways, same effects: Autonomous goal regulation is associated with subjective well-being during the post-school transition. *Motivation and Emotion*, *37*, 444–456.

Losier, G. F. (1994). Une analyse motivationnelle de la persévérance aux études avancées [A motivational analysis of graduate student persistence] (Unpublished doctoral dissertation). Université du Québec à Montréal, Montréal.

Lovitts, B. E. (2001). *Leaving the ivory tower: The causes and consequences of departure from doctoral study.* New York, NY: Rowman & Littlefield.

Lubke, G. H., & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling*, *11*, 514–534.

Marsh, H. W., Morin, A. J. S., Parker, P. D., & Kaur1, G. (2014). Exploratory structural equation modeling: An integration of the best features of exploratory and confirmatory factor analysis. *Annual Review of Clinical Psychology*, *10*, 85–110.

Marsh, H. W., Lüdtke, O., Muthén, B. O., Asparouhov, T., Morin, A. J. S., & Trautwein, U. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471–491.

Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., et al. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, *16*, 439–476.

McDonald, R. P. (1970). Theoretical foundations of principal factor analysis and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology*, *23*, 1–21.

McLachlan, S., Spray, C., & Hagger, M. S. (2011). The development of a scale measuring integrated regulation in exercise. *British Journal of Health Psychology*, *16*, 722–743.

Minister of Education, Recreation and Sports (2013). Indicateurs de l'éducation — Éd. 2012 [Education Indicators — 2012 Ed.]. Retrieved from http://www.mels.gouv.qc.ca/fileadmin/site_web/documents/PSG/statistiques_info_decisionnelle/Indicateurs_educ_2012_webP.pdf

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). *A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality.* Structural Equation Modeling (in press).

Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock, & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 395–436) (2nd ed.). Charlotte, NC: Information Age.

Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nettles, M. T., & Millett, C. M. (2006). *Three magic letters: Getting to Ph.D.* Baltimore, MD: Johns Hopkins University Press.

O'Meara, K., Knudsen, K., & Jones, J. (2013). The role of emotional competencies in faculty-doctoral student relationships. *The Review of Higher Education*, *36*, 315–347.

Organization for Economic Co-operation and Development (2013). Students enrolled by type of institution [statistic tables]. Retrieved from http://stats.oecd.org

Pauley, R., Cunningham, M., & Toth, P. (1999). Doctoral student attrition and retention: A study of a non-traditional Ed.D. program. *Journal of College Student Retention*, *1*, 225–238.

Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, *95*, 667–686.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioural research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., & Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*, 734–746.

Reamer, S. B. (1990). *Persistence of adult learners in an external doctoral degree program.* (Doctoral dissertation) Available from ProQuest dissertations and theses database (UMI No. 9023972).

Redsell, S. A., Lennon, M., Hastings, A. M., & Fraser, R. C. (2004). Devising and establishing the face and content validity of explicit criteria of consultation competence for UK secondary care nurses. *Nurse Education Today*, *24*, 180–187.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*, 749–761.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*, 68–78.

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wentzel, & A. Wigfield (Eds.), *Handbook on motivation at school* (pp. 171–196). New York, NY: Routledge.

Schmitz, J., Frenay, M., Neuville, S., Boudrenghien, G., Wertz, V., Noël, B., et al. (2010). Étude de trois facteurs clés pour comprendre la persévérance à l'université. *Revue Française de Pédagogie*, *172*, 43–61.

Sijtsma, K. (2009). On the use, misuse, and the very limited usefulness of Cronbach's alpha [introduction to a special issue]. *Psychometrika*, *74*, 107–120.

Statistics Canada and Human Resources and Skills Development Canada (2003). Access, persistence and financing: First results from the postsecondary education participation survey (catalogue no. 81-595M2003007). Retrieved from http://www.statcan.gc.ca/pub/81-595-m/81-595-m2003007-eng.pdf

Statistics Canada and Human Resources and Skills Development Canada (2009). Graduating in Canada: Profile, labour market outcomes and student debt of the class of 2005 (catalogue no. 81-595-M2009074). Retrieved from http://www.statcan.gc.ca/pub/81-595-m/81-595-m2009074-eng.pdf

Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, *80*, 99–103.

Vallerand, R. J., Blais, M. R., Brière, N. M., & Pelletier, L. G. (1989). Construction et validation de l'échelle de motivation en éducation (EME). [Construction and validation of the Motivation toward Education Scale]. *Canadian Journal of Behavioural Science/Revue Canadienne des Sciences du Comportement*, *21*, 323–349.

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, *72*, 1161–1176.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Sénécal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational & Psychological Measurement*, *52*, 1003–1017.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Sénécal, C., & Vallieres, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement*, *53*, 159–172.

Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K. M., & Deci, E. L. (2004). Motivating learning, performance, and persistence: The synergistic effects of intrinsic goal contents and autonomy-supportive contexts. *Journal of Personality and Social Psychology*, *87*, 246–260.

Vansteenkiste, M., Zhou, M., Lens, W., & Soenens, B. (2005). Experiences of autonomy and control among Chinese learners: Vitalizing or immobilizing? *Journal of Educational Psychology*, *97*, 468–483.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.

Wendler, C., Bridgeman, B., Cline, F., Millett, C., Rock, J., Bell, N., & McAllister, P. (2010). *The path forward: The future of graduate education in the United States.* Princeton, NJ: Educational Testing Service.

Wendler, C., Bridgeman, B., Markle, R., Cline, F., Bell, N., McAllister, P., & Kent, J. (2012). *Pathways through graduate school and into careers.* Princeton, NJ: Educational Testing Service.

West, S., Finch, J., & Curran, P. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes.* Los Angeles, CA: University of California.