Predicting stereotype endorsement and academic motivation in women in science programs: A longitudinal model

Marie-Noëlle Delisle  
School of psychology, Laval University, Quebec City, Canada

Frédéric Guay  
Faculty of Educational Sciences, Laval University, Quebec City, Canada

Caroline Senécal  
School of psychology, Laval University, Quebec City, Canada

Simon Larose  
Faculty of Educational Sciences, Laval University, Quebec City, Canada

Abstract

This study proposed and tested a model based on stereotype threat theory. The hypothesis is that women who are exposed to a low percentage of women in a science program are more likely to endorse the gender stereotype that science is a male domain, which will in turn undermine their autonomous academic motivation. A total of 167 women university students enrolled in science programs participated in an 18-month longitudinal study. Results partially support our model. Although the low percentage of females in science programs was related to endorsement of the gender stereotype, there was no effect of prior stereotype endorsement on subsequent autonomous academic motivation.

1. Introduction

In most industrialized countries, women make up the majority of the university student population. Although in mathematics programs men and women earn bachelor degrees in similar proportions, few women graduate in other scientific disciplines. For example, in 2003 in the province of Quebec, Canada, of the students that earned a bachelor degree in engineering, only 19% were women (CRSNG/ALCAN Chair, pour les femmes en sciences et génies au Québec, unpublished document, International Network of Women Engineers and Scientists INWES, 2004). Similarly, in the US in 2002, only 21% of engineering undergraduates were women (National Science Board, 2006). However, skilled and specialized workers in scientific fields are in high demand in the international market (NSF, 2004). Attracting more women to scientific fields and motivating them to their academic programs is a central challenge for our society.

Several studies have examined women's underrepresentation in science programs. Some have targeted the stereotype of science as a male domain (Aronson et al., 1998, Kahle and Meece, 1994, Seymour, 1995), because it can restrict women's course and career choices by projecting the image that these fields are inaccessible to women and that it is abnormal for them succeed in them (Kahle and Meece, 1994, Seymour, 1995).
In the present study, gender stereotype endorsement was examined in light of stereotype threat theory (Steele, 1997). We propose a two-step longitudinal model (see Fig. 1). In the first step, women who are exposed at the beginning of their science program (Time 1) to a low percentage of women in the program (i.e., low percentages of female students and teachers) are more likely to endorse the gender stereotype that science is a male domain at Time 1 (see Path A). Achievement was used as a control variable because previous studies have shown that achievement is related to both autonomous academic motivation and gender stereotype endorsement (see paths B and C in Fig. 1; Eccles, 1994, Guay and Vallerand, 1997, Ng and Bahr, 2000, Oswald and Harvey, 2003).

In the second step, women who endorse the gender stereotype at Time 1 are also hypothesized to endorse the gender stereotype at Time 2 (see Path D). However, endorsement of the gender stereotype at Time 1 is expected to be associated with lower autonomous academic motivation (acting out of choice and pleasure) at Time 2 (see Path E) when Time 1 autonomous academic motivation is taken into account (see Path F).

In the following sections, we provide empirical support for two crucial paths in the model, namely paths A and E. In fact, stability paths D and F are commonly identified in longitudinal studies (see Guay, Marsh, & Boivin, 2003), and paths B and C are not theoretically substantial, because achievement is used as a control variable rather than a theoretically explicative variable.

1.1. Gender stereotype endorsement and motivation (Path E)

1.1.1. Gender stereotype endorsement

Lippman (1922) described stereotypes as “the pictures in our head” that simplify the world by saving us the trouble of thinking too much when we come into contact with other people. These pictures are expectations of what people will be like and what they can and cannot do, and are usually generalized to all members of the group without considering individual differences (Bourhis, Gagnon, & Moïse, 1994).

Negative stereotypes have psychological consequences for those who are targeted. When individuals or groups are subjected to a negative stereotype, they will experience stereotype threat (Aronson, 2002, Croizet et al., 2003). According to stereotype threat theory (STT; Steele & Aronson, 1995), a threat refers to individual fears of confirming, through individual behaviors, a stereotype associated with the group. This threat interferes with individual performance, such that individuals unknowingly support the stereotype. For the threat to affect performance, three conditions must be present: 1) there must be a known negative stereotype about a social group in a particular field; 2) the individual must identify with that field and attach importance to it; and 3) the individual must be faced with a task that reveals the stereotype (Steele, 1997, Steele and Aronson, 1995). Our sample was composed of female science students. The three above conditions are met because 1) the women in our sample were subjected to the gender stereotype that science is more for men than women; 2) they attached some importance to their academic programs because they had selected them; and 3) they constantly faced academic tasks (exams, reports, etc.) that could reveal the stereotype.

Several STT studies have focused on the stereotype that women have less mathematical ability than men (Aronson and Steele, 2005, Smith and White, 2001, Spencer et al., 1999). Results of these studies show that women's performance is equivalent to that of men in situations where there is no obvious stereotype threat. However, when women are exposed to
a negative stereotype about their group (e.g., a difficult math test), they feel anxious about being judged along stereotypical lines and are under greater pressure to succeed. This greater pressure could lead to lower performance by women than men, which ironically confirms the stereotype (Martens et al., 2006, Smith and White, 2001, Spencer et al., 1999, Steele et al., 2002). In addition, studies on STT show that this threat is associated with other negative consequences, such as increased anxiety, decreased motivation toward the activity, and ultimately, disengagement (Croizet et al., 2001, Steele, 1997).

1.1.2. Autonomous academic motivation

We decided to focus on autonomous academic motivation as a consequence of stereotype threat. The concept of autonomous academic motivation stems from Self-Determination Theory (SDT, Deci & Ryan, 1985), and encompasses the reasons for doing an activity that relate to positive affective dimensions (pleasure, satisfaction) and identification with the activity (e.g., importance, utility) rather than doing the activity for external (e.g., rewards) or internal (e.g., guilt feelings) reasons. Autonomously motivated students achieve higher grades (Guay and Vallerand, 1997, Vansteenkiste et al., 2004) and are more persistent (Vallerand et al., 1997, Vansteenkiste et al., 2006).

SDT proposes that autonomous motivation flourishes when needs for competence (perceiving oneself as competent), relatedness (perceiving oneself as significantly connected to others), and autonomy (perceiving oneself as the instigator of one's own behavior) are met. Because gender stereotype endorsement conveys the message that sciences are more for men than for women, it threatens the three psychological needs, which in turn can decrease autonomous academic motivation. Consequently, we postulate that gender stereotype endorsement of science as a male domain may hamper women's autonomous academic motivation to pursue scientific studies (see Path E in Fig. 1).

1.2. Low percentage of females and gender stereotype endorsement (Path A)

For the threat to result in negative consequences, the stigmatized people must be immersed in a threatening environment (i.e., an environment that activates the stereotype). Various methods have been used to create a threatening environment, such as difficulty of the task or instructions, activation of gender stereotypes, and creation of disparaging female stereotypes through sexist television advertisements (Davis et al., 2002, Désert et al., 2002, Steele and Aronson, 1995). In this study, we posit that two features of science programs may produce higher gender stereotype endorsement, namely low percentages of 1) female students and 2) teachers.

1.2.1. Percentage of female students

Many studies indicate that women who are underrepresented in a group score lower on math tests and are more likely to think about gender stereotypes than women who take the test as part of a group with equal gender representation (Beaton et al., 2007, Inzlicht et al., 2006, Inzlicht and Ben-Zeev, 2000, Steele et al., 2002). Kanter, 1977, Kanter, 1989 explains that relationships between genders are influenced by their relative proportion in the group in a given environment. First, a particular group representing less than 20% of a group is in a token position (Kanter, 1977, Kanter, 1989). Women in a token position are usually isolated; it is difficult to form alliances, and relations with men are more difficult (Beaton and Tougas, 1997, Kanter, 1977, Kanter, 1989). Studies on token women show that they are viewed more stereotypically by their male colleagues and they feel that they have to work harder to prove
themselves (Beaton and Tougas, 1997, Kanter, 1977, Kanter, 1989, Whittock, 2002). Second, women making up around 35% of a group are in a minority position (Kanter, 1977, Kanter, 1989). Relations between men and women are still difficult, but the minority can form alliances among each other, which allows them to better adjust psychologically. Finally, when women make up about 50% of a group, the group is “balanced.” Balanced groups are characterized by harmonious inter-group relations, with more emphasis on personal characteristics than the characteristics of the subgroup to which one belongs (Kanter, 1977, Kanter, 1989).

In some science programs, such as the physical sciences and most engineering programs, women account for less than 20% of students and are in a token position. In other science programs, especially those related to biology and health, there is gender balance, with women accounting for 50% of students (CRSNG/Alcan Chair, unpublished document).

1.2.2. Percentage of female teachers

Statistics indicate that women make up less than 30% of teachers in science programs (CRSNG/Alcan Chair, unpublished document). Teachers are significant role models for students, and studies have shown that a lack of female role models or an overrepresentation of male role models can create a stereotype threat with consequent negative effects on women’s academic success and decisions on whether or not to work in the field (Bandura, 1997, Evans et al., 1995, Marx and Roman, 2002, McIntyre et al., 2005). In addition, several studies show that successful female role models can alleviate the gender stereotype effect (Marx and Roman, 2002, McIntyre et al., 2005). For example, self-efficacy and motivation toward non-traditional fields remain stable or increase in women who are exposed to successful female role models (Bandura, 1997, Marx and Roman, 2002, McIntyre et al., 2005).

Based on previous findings and the literature review, we hypothesized that the lower the percentages of female students and teachers in science programs, the more likely women students would be to endorse gender stereotype beliefs at Time 1 (see Path A, Fig. 1).

1.3. The present study

Our study brings a unique contribution to the literature. First, most studies on STT used samples of psychology students to evaluate the effects of stereotype threat on women's performance. In contrast, our proposed model was tested on a sample of women enrolled in science programs. This increases the generalization of results, because the women in our sample were directly targeted by the stereotype that they did not have the required abilities. Second, whereas most studies on STT used an experimental design, the proposed model was tested in a natural setting (Inzlicht and Ben-Zeev, 2003, Kray et al., 2001, Sax, 1996), thereby increasing the ecological validity of the findings. Third, in contrast to past studies that use a cross-sectional design, our study used a longitudinal design, a particularly powerful approach to test the model outlined above. Finally, the percentages of female students and teachers in science programs were evaluated using objective criteria. Consequently, if these variables correlated significantly with gender stereotype endorsement, we could reasonably conclude that the low percentage of females in science programs caused gender stereotype endorsement, and not the contrary.
2. Method

2.1. Sample and procedure

Participants were women university students enrolled in various undergraduate science programs at Laval University, a large French-speaking university in Quebec City, Canada. During the first week of the autumn semester (September 2003, T1), research assistants visited the students in their classes and asked them to fill out a questionnaire at home. A total of 1431 students entered science programs at Laval University in 2003 (part-time and full-time). Of these, 1028 were men and 403 were women. A total of 167 women students filled out a questionnaire for a response rate of 41%.

All students were in their first semester at university and the mean age was 19.74 years. The data show that 66% of their mothers worked full time (other mothers worked part time or were housewives), whereas 87% of their fathers worked full time (other fathers were retired or worked part time). Income varied according to the parents' gender: mean annual income was $41,410 for fathers and $30,164 for mothers. A total of 56 students lived with their parents during the university semester (34.15%).

In March 2005 (T2), 18 months later, the students were contacted by phone and invited to participate in the second part of the study. Students were sent a questionnaire by mail and asked to fill it out. A total of 148 participants returned their questionnaires, for an 88% response rate.

2.2. Measures

2.2.1. Percentages of women students and teachers

Percentages of women students and teachers in each program were determined using official statistics published by Laval University (Bureau du registraire de l'Université Laval, 2004). Percentages of women in different programs ranged from 10% to 69%, whereas percentages of women teachers ranged from 0% to 27%. On each of these two variables, each student was scored on the percentage of women teachers and students in a given program. For example, if a student was studying biology, the percentages of female students and female teachers in that discipline were entered separately for that student. The correlation between the two indicators was 0.52, which supports their use as two indicators of the latent construct “low percentage of females in science programs” (see Fig. 1).

2.2.2. Gender stereotype endorsement scale

This scale is based on the Fennema–Sherman Mathematics Attitude Scales (FSMAS, Mulhern & Rae, 1998). It was administered at both measurement times (T1 and T2). The FSMAS is a multidimensional scale containing 108 statements divided into nine scales measuring a variety of attitudes toward mathematics. In this study, only the “mathematics as a male domain” subscale was used to evaluate gender stereotype endorsement. This subscale specifically assesses the stereotype that women cannot study mathematics as well as men do and that mathematics is considered more suitable, important, and interesting for men than women. The subscale includes 12 items scored on a 5-point Likert scale (“strongly disagree” to “strongly agree”). The Fennema–Sherman scale has been used in many studies and has shown excellent psychometric qualities (Broadbooks et al., 1981, Hyde et al., 1990, Mulhern and Rae, 1998, Tocci and Engelhard, 1991). Moreover, studies that adapted this scale to
measure attitudes toward science also revealed very good psychometric properties (Bazler et al., 1993, Collier et al., 1998, Levin and Fowler, 1984).

In this study, nine items of the 12-item subscale were adapted to assess stereotypes in science simply by changing the word “mathematics” to “science” in each of the nine statements. One item, “Mathematics is for men; arithmetic is for women,” was reworded to, “Science and engineering programs are for men; the humanities are for women” for greater consistency with the study goal. Note that this nine-item version has been proven valid in past studies. For example, Mulhern and Ray (1998) showed that loadings on this short version ranged from 0.42 to 0.70 with an alpha reliability coefficient of 0.85. In our study, internal consistency values at the two measurement times were 0.70 (Time 1) and 0.76 (Time 2).

2.2.3. Autonomous academic motivation

The Academic Motivation Scale (AMS) was used to assess autonomous academic motivation (Vallerand, Blais, Brière, & Pelletier, 1989). The AMS was administered at both measurement times, and was adapted to assess student motivation toward science programs. Thus, each item represents a possible reason for studying science. The original version of this scale includes 28 items subdivided into seven subscales (Vallerand et al., 1989). In the present study, we used only five subscales of the AMS: intrinsic motivation toward knowledge (Time 1 mean = 5.95, Time 2 mean = 5.72, “I study science because I get pleasure and satisfaction out of learning new things”), identified regulation (Time 1 mean = 5.75, Time 2 mean = 5.49; “I study science because eventually it will allow me to enter the job market in a field that I like”), introjected regulation (Time 1 mean = 3.33, Time 2 mean = 3.27 “I study science to prove to myself that I can succeed in this field of studies”), external regulation (Time 1 mean = 3.44, Time 2 mean = 3.30 “I study science because this field of studies will enable me to find a paying job”), and a motivation (Time 1 mean = 1.38, Time 2 mean = 1.77; “Honestly, I don't know; I really feel that I'm wasting my time in this program”). The other two intrinsic motivation subscales were not used because they are strongly correlated with intrinsic motivation to knowledge. For example, in Vallerand et al. (1993), intrinsic motivation to knowledge correlated at 0.58 with intrinsic motivation to accomplishment, whereas intrinsic motivation to stimulation correlated at 0.59 with intrinsic motivation to knowledge, and finally, intrinsic motivation to stimulation correlated at 0.62 with intrinsic motivation to accomplishment (see also Vallerand et al., 1989 for similar correlation coefficients). These high correlations were not observed for the three types of extrinsic motivation, however. For example, correlations between types of extrinsic motivation ranged from 0.29 to 0.48 in Vallerand et al. (1993) and from 0.30 to 0.36 in Vallerand et al. (1989), which supports their inclusion in this study. Items were scored on a 7-point Likert scale (“do not correspond at all” to “correspond strongly”). A preliminary analysis was carried out to evaluate the factorial validity of the AMS. Results indicated that the factor solution involving the five subscales was not optimal at Time 1 ($\chi^2[160, N = 167] = 354.98, p < .05, \chi^2/df = 2.22, \text{RMSEA} = .086$) or Time 2 ($\chi^2[160, N = 125] = 308.31, p = .05, \chi^2/df = 1.93, \text{RMSEA} = .081$). Despite these non-optimal fit indices, most factor loadings were substantial. Chronbach's alpha values for these subscales were also adequate, ranging from 0.61 to 0.88 for Time 1 and from 0.77 to 0.90 for Time 2.

The five subscales were used to compute a relative autonomy index (RAI, Grolnick & Ryan, 1989) to reduce the number of variables by integrating the score for each motivation subscale into a single score. Thus, a positive weight was given to autonomous items and a negative weight to controlled items, according to their position on the autonomy continuum (Fortier, Vallerand, & Guay, 1995). To form the RAI, the amotivation subscale was weighted −2, the
external regulation subscale and the introjected subscale were weighted − 1, the identified subscale was weighted + 1, and the intrinsic subscale was weighted + 2. In other words, the more controlled the regulatory style in a subscale, the greater its negative weight, and the more autonomous the regulatory style in a subscale, the greater its positive weight. Index scores fluctuated from − 18 to + 18, where a high score reflects high autonomous academic motivation. Several studies have shown the usefulness of the RAI (Black and Deci, 2000, Deci and Ryan, 2000, Grolnick and Ryan, 1987).

2.2.4. Achievement

In order to control for student ability, we used student achievement in college. College education in the Quebec educational system refers to a post-high school but pre-university system offering 2-year (pre-university) and 3-year (terminal technical) programs. Achievement was assessed using an R score. Full computation of the R score is not presented here due to space constraints. Briefly, the R score is a standardized measure of achievement used in the Quebec education system to select students for admission into various university programs, among others. Average R score for the entire Quebec student population is approximately 25 (CREPUQ, 2004).

2.3. Statistical analyses

2.3.1. Missing data estimation

It would be highly inappropriate to use a Listwise deletion of cases for the missing data. Several researchers have shown that this method as well as other ad hoc methods, such as substituting the variable mean for missing values, are inappropriate for treating missing values (Davey et al., 2001, Peugh and Enders, 2004). In the present study, a full information maximum likelihood (FIML) approach was used to estimate missing values. Briefly, the covariance matrix and sample mean estimates were rebuilt. Many studies suggest that this method generally produces the least biased and most efficient parameter estimates (Jamshidian and Bentler, 1999, Peugh and Enders, 2004). Thus, all analyses described in the Results section were based on a sample of 167 participants for whom missing values at Time 2 (n = 19) were estimated.

2.3.2. Goodness of fit indices

Model adequacy was assessed by structural equation modeling (SEM) using LISREL (Version 8.80). All models were tested with maximum likelihood estimation. To ascertain model fit, we used the root-mean-square error of approximation (RMSEA) and the chi-square/degrees of freedom ratio ($\chi^2/df$). According to Browne and Cudeck (1993), RMSEA values less than 0.05 are considered a good fit, values from 0.05 to 0.08 provide an adequate fit, and values from .08 to 0.10 provide a mediocre fit. Values > 0.10 are unacceptable. The chi-square/degrees of freedom ($\chi^2/df$) ratio is a function of model misfit ($\chi^2$) compared to model parsimony, as indicated by the model's degrees of freedom (df). Smaller $\chi^2/df$ ratios occur when model misfit is lower than model parsimony. In general, a $\chi^2/df$ ratio less than 2 indicates a relatively good model fit (Kline, 2005). The CFI and NNFI indices were not used because these fit indices are not produced under FIML in LISREL.

2.3.3. Correlated uniquenesses and correlated disturbance

Covariances among uniquenesses were estimated in the models tested. When the same items are administered to the same participants on multiple occasions, it is likely that the
uniquenesses associated with matching items will be correlated. If these correlated uniquenesses are not estimated in the model, the estimated correlation between the corresponding latent constructs will be positively biased (Guay et al., 2003, Marsh and Yeung, 1998). In other words, they could lead to systematically inflated estimates of stability. It is also recommended to account for this potential source of errors when the same items are used on multiple occasions (Guay et al., 2003, Marsh and Yeung, 1998). Consequently, we estimated the correlated uniquenesses in our confirmatory and structural models. Furthermore, the covariance between disturbance terms was estimated.

2.3.4. Invariance of factor loading

Factor loadings across measurement times were constrained to equality in our models to ensure that participants understood the constructs in the same way over time. This method ensures that changes in the amount of reliable variance indicators are adequately captured as changes over time in the common construct variance (Little, Preacher, Selig, & Card, 2007). In this study, this means that observed effects on autonomous academic motivation and gender stereotype endorsement are not confounded with changes in the meaning of the construct.

2.3.5. Transformation of scores

Because some variables were not normally distributed, we computed normal scores for all variables from the ranks using Blom’s formula (Blom, 1958): \( y_i = \Phi^{-1}(r_i - 3/8)/(n + 1/4) \), where \( \Phi^{-1} \) is the inverse cumulative normal (PROBIT) function, \( r_i \) is the rank of the \( i \)th observation, and \( n \) is the number of non-missing observations for the ranking variable.

3. Results

We began with a CFA analysis to estimate correlations among the constructs, and descriptive statistics for all variables (means and standard deviations) are provided below. Finally, we tested our longitudinal model.

3.1. Descriptive statistics and CFA correlations

Means and standard deviations are presented in Table 1. Results indicated that, at both measurement times, few women endorsed the gender stereotype that science is a male domain. Moreover, at Time 1 and Time 2, most women had autonomous reasons to pursue their science program. Repeated measure analyses indicated that gender stereotype endorsement \( (F (1,162) = 0.59, p = 0.4435) \) did not vary over time. However, autonomous academic motivation \( (F (1,162) = 16.21, p < .05) \) significantly decreased from Time 1 to Time 2.

A CFA analysis was conducted to explore correlations among variables, which obtained good fit indices \( (\chi^2 [103, N = 167] = 127.92, p = .05, \chi^2/df = 1.24, \text{RMSEA} = .0.38 \in [.003; .058]) \). Results indicated stability correlations for autonomous academic motivation and gender stereotype endorsement at 0.50 and 0.62, respectively (see Table 1). Consistent with our model, low percentage of women in science programs was positively related to gender stereotype endorsement at Time 1 \( (r = 0.30) \) and Time 2 \( (r = 0.28) \). The expected negative correlation between Time 1 gender stereotype endorsement and Time 2 autonomous academic motivation was not significant, although Time 1 autonomous academic motivation was negatively correlated to Time 2 gender stereotype endorsement \( (r = - .38) \).
3.2. SEM

In line with our hypotheses, we tested the SEM model shown in Fig. 1. For greater stringency, we estimated two additional paths (not shown) connecting low percentage of females to Time 1 autonomous academic motivation and connecting Time 1 autonomous academic motivation to Time 2 gender stereotype endorsement (see Marsh, Byrne, & Yeung, 1999). Fit indices were adequate ($\chi^2 [108, N = 167] = 132.24, p = .06, \frac{\chi^2}{df} = 1.22$, RMSEA = .040 $\in [0.00; 0.06]$). Results showed (Fig. 2) that Time 1 low percentage of females predicted Time 1 gender stereotype endorsement ($\beta = .32$). Moreover, Time 1 autonomous academic motivation positively predicted Time 2 autonomous academic motivation ($\beta = .49$), and Time 1 gender stereotype endorsement positively predicted Time 2 gender stereotype endorsement ($\beta = .59$). In contrast to our hypothesis, Time 1 gender stereotype endorsement did not significantly predict Time 2 autonomous academic motivation, but Time 1 autonomous academic motivation significantly predicted Time 2 gender stereotype endorsement ($\beta = -.27$). These results call into question the second step of our proposed model.

4. Discussion

The goal of the present study was to test a model that posits that when women account for a low percentage in a science program (assessed by the percentages of female students and teachers in science programs), they are more likely to endorse the gender stereotype that science is a male domain. In addition, women who get good grades in college are expected to endorse the gender stereotype to a lesser degree, but they should self-report higher autonomous academic motivation. In the second step, women who endorse gender stereotype are hypothesized to have lower subsequent academic motivation (acting out of choice and pleasure).

Results based on CFA correlations showed that the more women were exposed to a low percentage of females in their science programs, the more likely they were to endorse the stereotype of science as a male domain. Additionally, findings indicated that women who were autonomously motivated endorsed this gender stereotype less. Longitudinal SEM analyses showed similar findings. Specifically, our results indicated that a low percentage of females in science programs was associated with higher gender stereotype endorsement. However, there was no significant path connecting prior gender stereotype endorsement to subsequent autonomous academic motivation, which casts doubt on the mediation role of gender stereotype endorsement. However, prior autonomous academic motivation predicted subsequent gender stereotype endorsement. These results are discussed in more detail in the following sections.

4.1. Low percentage of females in science programs and gender stereotype endorsement

Results of the longitudinal SEM model indicated that when women students accounted for a low percentage in their science programs (i.e., the number of students and teachers), they showed higher gender stereotype endorsement at the beginning of the program. In other words, a minority position can be sufficient to produce unfavorable consequences for women, probably because it fosters undue attention to group characteristics and leads to excessive concerns about the stereotype that science is for men and not for women (Inzlicht et al., 2006, Inzlicht and Ben-Zeev, 2000, Inzlicht and Good, 2006, Marx and Roman, 2002, McIntyre et al., 2005).
It should be noted that the relationship between a lower percentage of females and gender stereotype endorsement was significant even when controlling for students' initial competencies. Thus, low exposure to females in science programs may explain a portion of the variance that cannot be explained by achievement. However, results indicated that college achievement related to neither gender stereotype endorsement nor women's autonomous academic motivation. This is somewhat surprising, in light of past research (Eccles, 1994, Ng and Bahr, 2000, Oswald and Harvey, 2003). One possible explanation is that achievement was assessed in college and not during university semesters. Using a measure of university achievement may lead to positive correlations among these constructs in future research.

4.2. Gender stereotype endorsement and autonomous academic motivation

The longitudinal data analysis showed that gender stereotype endorsement at Time 1 was not associated with autonomous academic motivation at Time 2. However, other variables may have played a mediating or moderating role in the relationship between gender stereotype endorsement and autonomous academic motivation. For example, many studies have identified different mediators between stereotype threat and test performance, such as anxiety, expectations, and cognitive load (Aronson and Steele, 2005, Désert et al., 2002). This suggests that negative stereotypes impair performance by depleting cognitive resources. Hence, future research on these mediators may explain decrease in autonomous academic motivation. Another possible explanation is that achievement was assessed in college and not during university semesters. Using a measure of university achievement may lead to positive correlations among these constructs in future research.

Moreover, the type of measure used to assess gender stereotype endorsement may explain some of the non-significant results. Due to cultural and/or personal pressures, people tend to conceal their negative attitudes toward different groups. This interpretation is consistent with our results that showed low means at both measurement times for gender stereotype endorsement. To remedy this bias, researchers have suggested assessing people's attitudes implicitly (Greenwald et al., 1998, Nosek et al., 2002). Automatic, unconscious attitudes can be measured with implicit tests, as there is no opportunity to control responses based on what is culturally acceptable. Recent studies on implicit measures (Nosek et al., 2002) showed that students endorse stereotype more implicitly than explicitly. Thus, even though women may state that they do not believe in a gender stereotype, the fact that they know about the stereotype or endorse it implicitly is enough to have negative effects on performance and motivation (Aronson & Steele, 2005).

It is important to note, however, that prior autonomous academic motivation negatively predicted subsequent gender stereotype endorsement. This finding could be explained by the possibility that students who score high on autonomous motivation have a strong internal need to learn. Consequently, they may be less likely to be affected by contextual factors such as lack of women classmates and teachers.
Helping students to value and get pleasure out of their science programs therefore appears to be a promising avenue to reduce stereotype endorsement. This increase in autonomous academic motivation could be achieved through autonomy-supportive practices, in which the teacher considers the students' perspectives, provides a rationale for requests, acknowledges their feelings and perceptions, provides them with information and choices, and minimizes the use of pressure and control such as task deadlines, performance-based rewards, imposed goals for a given activity, and competition (Ryan & Deci, 2000). Some studies (Vallerand et al., 1997) have shown that autonomy-supportive teachers tend to have more autonomously motivated students.

4.3. Limitations

This study has a number of shortcomings. First, although a low percentage of female students in science programs was linked to both gender stereotype endorsement and autonomous academic motivation, we cannot definitively conclude that this construct was the real cause. That is, specific characteristics (other than those evaluated) of the programs may have influenced gender stereotype endorsement and autonomous academic motivation. For example, professors' expectations and teaching methods, level of academic difficulty, and the use of competitive learning activities can vary between programs (Cronin and Roger, 1999, Seymour, 1995, Young and Fraser, 1994).

Second, it is arguable that, because the students in this study were initially assessed during the first week of the autumn semester, it is possible that gender stereotype threat had been activated before they entered university, i.e. in college where female students could also have been exposed to few women. However, it is important to note that, in 2003 (when this study was conducted), the percentages of women and men registered in college science programs were roughly equivalent (50% women; Chesco, 2008). It therefore appears that the percentage of women that we found in university science programs affected gender stereotype endorsement, not their previous experience with low ratios in college.

Third, some non-significant effects were most probably due to the small sample size, which was insufficient to capture the small effects observed in the model. Fourth, fit indices stemming from the factor structure of the Academic Motivation Scale (AMS; Vallerand et al., 1989) were somewhat low, especially at Time 1. Nevertheless, factor loadings were appropriate, as well as reliability coefficients. The low fit indices may be explained by the small sample size. For example, using a larger sample, Grouzet, Otis, and Pelltier (2006) obtained higher fit indices for the AMS.

5. Conclusion

Despite these limitations, it is important to note that our results support the ecological validity of previous studies, because they were derived from a natural environment. Compared with experimental studies, in which students are placed in an exceptional context, the present study used female students who attended classes in which the gender stereotype threat could be activated. Based on our results, one way to reduce these negative consequences would be to attract more women professors to scientific programs and to increase the number of women attending science classes. However, this is easy to say but not so easy to achieve. In future research, it would be useful to investigate factors that could moderate the effects of low percentage of females in science programs on gender stereotype endorsement. Nevertheless, as
explained earlier, a good way to reduce gender stereotype endorsement might be to increase autonomous academic motivation. To this end, the use of autonomy-supportive practices would be a promising avenue (Ryan & Deci, 2000).
References


Figure 1. The proposed longitudinal model.

Step 1
- Time 1: The Beginning of the Science Programs, September-2003

Step 2
- Time-2: 18-months Later in the Science Programs, March-2005
Figure 2. Results of the proposed model.
Table 1. CFA correlations among variables.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum/Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low percentage of females–Time 1a</td>
<td>–</td>
<td>0.30*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>34.95</td>
<td>13.30</td>
<td>20.13/58.05</td>
<td></td>
</tr>
<tr>
<td>Gender stereotype endorsement–Time 1</td>
<td>0.30*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.15</td>
<td>0.28</td>
<td>1.00/2.89</td>
<td></td>
</tr>
<tr>
<td>Gender stereotype endorsement–Time 2</td>
<td>0.28*</td>
<td>0.62*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.13</td>
<td>0.28</td>
<td>1.00/2.78</td>
<td></td>
</tr>
<tr>
<td>Autonomous academic motivation–Time 1</td>
<td>–0.15</td>
<td>–0.19</td>
<td>–0.38*</td>
<td>–</td>
<td>–</td>
<td>11.52</td>
<td>3.26</td>
<td>–0.63/17.50</td>
<td></td>
</tr>
<tr>
<td>Autonomous academic motivation–Time 2</td>
<td>–0.10</td>
<td>–0.16</td>
<td>–0.13</td>
<td>0.50*</td>
<td>–</td>
<td>10.10</td>
<td>4.67</td>
<td>–9.13/16.50</td>
<td></td>
</tr>
<tr>
<td>Achievement in college</td>
<td>0.09</td>
<td>–0.09</td>
<td>–0.19</td>
<td>0.04</td>
<td>0.14</td>
<td>–</td>
<td>29.81</td>
<td>3.78</td>
<td>18.52/37.40</td>
</tr>
</tbody>
</table>

*a p < 0.05.

Averages of scores on low percentage of 1) female students and 2) female teachers.